FACULTY OF PSYCHOLOGY
AND EDUCATIONAL SCIENCES

Binding Benefits in Visual Working Memory: Investigating the Role of Spatial Attention in Feature Integration

Word count: 11.803

Elisa Verhofstadt
Student number: 01703533

Promotor: Prof. Dr. Wim Fias
Co-promotor: Dr. Muhammet Ikbal Sahan

A dissertation submitted to Ghent University in partial fulfilment of the requirements for the degree of Master of Science in Psychology (Theoretical and Experimental Psychology)

Academic year: 2021 – 2022

GHENT
UNIVERSITY

Acknowledgments

Throughout the process of writing this Master thesis I have enjoyed the support of many wonderful people, without whom the completion of this thesis would not be possible. First of all, I would like to thank prof. dr. Wim Fias for the opportunity to execute this study and for the introduction to this interesting field of research. Secondly, I would like to thank dr. Muhammet Ikbal Sahan for the valuable guidance throughout every step of this project. His encouragement, feedback and expertise greatly helped to improve the quality of this thesis. Also, my gratitude goes out to my parents for providing me this opportunity and continuously supporting me throughout these years. Lastly, I want to express my gratitude towards Jens for his unconditional support and genuine interest in my research topic.

Abstract

The ability to integrate several features into one object greatly increases the capacity of visual working memory (VWM). This finding has been taken as evidence that attention in VWM operates in an object-based manner. However, these classical experiments of feature binding are unable to disentangle object-based attention from spatial attention in VWM. Since features of an inherent object inevitably share the same location, an alternative explanation of feature integration through shared spatial attention is proposed. This study aimed to test this spatial binding hypothesis by presenting two inherent random dot kinematograms, with dots containing both color and direction features (bound samples), and comparing them to two non-inherent random dot kinematograms, with dots containing either a color or direction feature (unbound samples). Precision of VWM representations was assessed by means of a continuous reproduction task which required participants to recall one color and one direction feature. Retro-cues were used to manipulate spatial attention by selecting the to-be-recalled features from either the same sample (spatially overlapping condition) or from two different samples (spatially non-overlapping condition). The results indicated a similar recall precision for both bound and unbound samples, supporting the spatial binding hypothesis. However, this equivalence could only be found during the spatially overlapping condition, while the spatially non-overlapping condition showed an advantage of bound sample recall. This pattern of results suggests a dissociation between two types of features binding, (1) binding through shared location when spatial attention is available, and (2) direct feature binding requiring no attention. The implications of this dichotomy are discussed further.

Nederlandstalige samenvatting

Het vermogen om verschillende kenmerken in één object te integreren verhoogt de capaciteit van ons visueel werkgeheugen (VWG) enorm. Deze bevinding wordt als bewijs aangevoerd voor het feit dat aandacht in het VWG op een object-gebaseerde manier werkt. Echter, deze klassieke experimenten kunnen onmogelijk object-gebaseerde aandacht van ruimtelijke aandacht in het VWG onderscheiden. Omdat kenmerken van inherente objecten onvermijdelijk dezelfde locatie delen, wordt er een nieuwe verklaring voorgesteld waarbij kenmerken geïntegreerd worden door beroep te doen op gedeelde ruimtelijke aandacht. Deze studie beoogt deze ruimtelijke bindingshypothese te testen door twee inherente ‘random dot kinematograms’, waarbij stippen zowel een kleur als richting kenmerk bevatten (gebonden sample), te vergelijken met niet-inherente ‘random dot kinematograms’, waarbij stippen enkel een kleur of richting kenmerk bevatten (ongebonden sample). De precisie van de VWG representaties werd beoordeeld door middel van een continue reproductietaak, waarbij participanten telkens één kleur en één richting moeten oproepen. Retro-aanwijzingen werden gebruikt om ruimtelijke aandacht te manipuleren door de op te roepen kenmerken te selecteren van ofwel dezelfde sample (ruimtelijk overlappende conditie) of van verschillende samples (niet-ruimtelijk overlappende conditie). De resultaten toonden een gelijkaardige precisie aan voor zowel gebonden als ongebonden samples, wat de ruimtelijke bindingshypothese ondersteunt. Echter, deze gelijkheid kon enkel gevonden worden bij de ruimtelijke overlappende conditie, terwijl de niet-ruimtelijk overlappende conditie een voordeel toonde van gebonden samples op te roepen. Dit patroon van resultaten suggereert een dissociatie tussen twee vormen van kenmerkintegratie, (1) integratie door gedeelde locatie wanneer ruimtelijke aandacht beschikbaar is, en (2) directe kenmerkintegratie zonder aandachtsbenodigdheden. De implicaties van deze tweedeling worden verder besproken.

Table of contents

Introduction

When looking at a scene, even for just a slight moment, humans do not require much effort to perceive a coherent picture of their environment. You easily recognize grass as being green, a ball on this grass as having a round shape and a red color, while you register a blue butterfly flying across the garden in a particular direction. However, this is quite an impressive ability as you consider that the brain processes all aspects of color, shape and motion simultaneously through separate early visual systems. In this way, it is extremely important that the brain binds these independent features again into their respected objects. Otherwise, one might end up with a scene where the grass is blue with a red butterfly flying across. Substantial research has looked into how this binding is possible (the binding problem) and how objects are processed by our perceptual system. Importantly, there is evidence that the attentional mechanisms involved in visual perception are also important in guiding attention in visual working memory (Awh et al., 2006; Chun, 2011; Chun et al., 2011; Gazzaley & Nobre, 2012; Postle, 2006). Likewise, we see that even in visual working memory the ability to remember integrated objects is present (Allen et al., 2006; Delvenne & Raymond, 2004; Luria & Vogel, 2010; Sahan et al., 2019; Xu 2002). For example, when we are trying to remember the ball we had just seen on the grass, we do not simply retrieve' something red' and 'something round'. In fact, we retrieve an integrated object, which is simultaneously round and red, from our visual working memory. Because our visual working memory only has limited capacity, it is beneficial to store items as integrated objects (a round and red ball) rather than their separate features. Despite this, it is still unclear how binding in visual working memory is exactly achieved. More specifically, it is possible that these features are stored together through their shared spatial location, rather than being bound into an independent integrated object. In this master thesis, the influence of spatial overlap for object-benefits in visual working memory will be investigated.

Perceptual Features in Vision

The perceptual properties of features have often been studied using a visual search paradigm. In this paradigm, participants are instructed to identify whether a target stimulus is present in a display that also contains distractors. Target stimuli are defined by the presence of (a) specific feature(s). Often, the relation between reaction times (RT) and set size is investigated, as this might tell us something about how the attentional system operates. By varying the features of the target stimulus and distractor items, researchers try to determine how these features guide our visual search, or in other words, how our visual processing and goal-directed attention is influenced by features (Egeth et al., 1984; Nakayama & Silverman, 1886; Theeuwens, 1996; Theeuwens & Kooi, 1994; Wolfe et al., 1989). For instance, Wolfe and Horowitz (2017) defined a feature as a visual property that directly guides our attention during visual search, independent of set size. Therefore, this creates a 'pop-out' effect in the search display. They illustrated this by showing a display set of stimuli with T-intersections and observing which targets facilitated the pop-out effect. Stimuli with a different color compared to their neighbors and stimuli with a different orientation all showed this pop-out effect. In other words, targets get recognized immediately no matter how big the set size. However, stimuli whose intersection pattern made them look like a cross (X-intersection target among T-intersection distractors), do not guide our attention in the same way. Recognition of these stimuli takes longer because all the neighboring stimuli also contain similar intersections. Therefore, recognition time would be dependent upon the set size, which means that the recognition time would further increase as larger set sizes are shown. This initial visual search research therefore defined serial and parallel attentional searches. Serial search is highly dependent on the set size, because attention has to be deployed to every item location individually before it can be rejected or accepted. In parallel search, feature salience guides bottom-up attention in such a way that the display size is irrelevant. For example, Egeth (1984) showed participants a display with black O- and red N-letter distractors accompanied by a single red O as target letter. RT for identification of the red O target was highly correlated to the set size of its distractors. However, when researchers specifically inform participants this target feature is red, the RT x set size slope is reduced by 50%. This can be explained by parallel search of the color stimuli. Importantly, color can automatically guide attention in visual search tasks (Wolfe & Horwitz, 2004). In this way, participants are able to directly eliminate all non-target colors in this display. Thus, independently of the set size, search time will always be cut in half when color information is provided. On the other hand, search for the O target happened in a more serial fashion. Letter identity is not perceived automatically (Wolfe & Horwitz, 2004). In this case, the set size, now reduced to the amount of letters with the target color, does correspond to search time. Attention has to be devoted to each individual letter to determine whether it corresponds to the target letter. In conclusion, visual search paradigms can help us distinguish between the influence of different types of features on selective visual attention. Some features like color, orientation, size, depth, motion (Wolfe & Horwitz, 2004) can directly capture our attention and make use of parallel search.

Importantly, visual search experiments show that objects defined by a conjunction of features usually display serial search, while objects defined by a single feature display parallel search (Eckstein et al. 2000; Woodman & Luck, 2003). Treisman (1982) was one of the first to show that targets defined by a conjunction (e.g., a green T-shape) showed a steep RT x set size slope. In contrast, targets defined by single features (e.g., the color green) which show a flat slope as set size increases. This discovery led to the development of Feature Integration Theory (FIT), a two-stage theory that explains the attentional mechanisms behind the detection of features and their conjunctions during perception (Treisman & Galade, 1980). The theory states that individual features are represented in a feature map across the visual field. This map reflects the representation of all features with their corresponding locations. Furthermore, these maps can be created and consulted without deploying direct attention, as they are formed when first observing a scene. Several feature maps are created simultaneously for each of the feature dimensions present in any context (color feature maps, shape feature maps,…). During a visual search task, when the target item is defined by a single feature, the result of this automatic feature map access can be observed as the pop-out effect. Because of the possibility of pre-attentive processing of these one-dimensional features, the location of a target feature can be found without any additional guidance of attention in the perceptual space. In such a case, participants show flat search slopes during visual search. Therefore, this indicates parallel processing of these one-dimensional features.

In contrast, when a target stimulus is defined as a conjunction of several features, attention needs to be deployed to bind these features into an integrated object. More specifically attention to the object’s location is necessary according to FIT. Before attention is deployed, these same features can be interpreted as ‘free-floating’ in their respected feature maps. For instance, when participants are instructed to search for a blue triangle among distractors they have to allocate their attention to each location of the display during this task. In this case, the individual ‘blue’ feature map pre-attentively indicates all the locations where blue features are present. Similarly, the ‘triangle’ feature map does the same for all triangle stimuli. However, these feature maps need to be combined to detect the blue triangle. The combination of these features is only available once attention is allocated to the location of an item. This causes the ‘blue’ and ‘triangle’ feature map to merge for that particular location. As a result, the object with both of these features, a blue triangle, is perceived. So, when participants have the goal of detecting this object, the focus of attention should be shifted from one location to another in order to identify the target. Consequently, serial search happens during the search for multi-feature objects. Accordingly, slower RTs are observed in larger set sizes during the search for these multi-feature objects.

It should be noted that FIT’s assumptions about serial and parallel search has been criticized because this dichotomy cannot be recreated within more diverse situations (de-Wit, 2011; Liesfeld & Müller, 2019). Conjunction search can sometimes facilitate a pop-out effect, while feature search sometimes takes place serially. Likewise, more evidence points towards a continuum between serial and parallel search (Duncan & Humphreys 1989; Liesefeld & Müller, 2019; Wolf, 1998) rather than two fully distinguished systems. Other theories of visual search each emphasise other mechanisms by which attention is allocated (Wolfe, 2020). Guided Search (GS) of Wolfe (1989) claims that parallel processed feature information can guide focal attention as to what location to attend next. Attention is not only guided by bottom-up salience and top-down target features, but additionally by scene, search history, and value (Wolfe & Horowitz, 2017). These features are combined into priority maps, which select relevant locations according to the point with the highest activity. The dimensional weighing model (Liesefeld & Müller, 2019) expands upon this idea by claiming that different features have the ability to guide attention in different ways dependent on task requirements. Furthermore, the Target Acquisition Model (TAM) (Zelinsky, 2008) proposes that parallel search happens in ‘clumps’ created by a point of fixations. This is in contrast to FIT, which proposes that only one location can be attended to at the same time. Clumps in the TAM are thought to be able to include several stimuli simultaneously. These clumps must be processed in a serial way to comprehend the whole display.

Altogether, theories of visual search are often still unclear about the specific mechanism behind the allocation of attention to objects, as well as under which circumstances this attention can be allocated serially or in parallel. Nonetheless, most of these theories seem to agree with some general statements about the allocation of attention to objects. More specifically, due to the limited capacity of attentional resources, search for more complex objects made up of multiple features must contain some aspect of serial search. It is impossible to acquire full representations of every single object in a display or context without using focal attention. For example, Wolfe and Bennett (1997) showed participants displays with crosses containing a vertical and a horizontal bar which could be either red or green. Here, a target item is always the binding of some basic features (e.g., a vertical green bar together with a horizontal red bar), and thus complex. RT on this task was heavily dependent upon set size. Before attention was allocated, all items in the display were 'pre-attentive' and thus unable to be identified as a bound object. Attention is needed to fully represent a complex item.

Feature Binding in Perception

The literature on visual search introduces a very interesting question, how can separate features be combined into one object representation? This is what is called the 'binding problem'. This question has also been asked outside the field of visual search research and it is considered a very important problem in neuroscience (Burwick, 2014). Von Malsburg (1999) was the first to name the binding problem and illustrated how difficult it is for neural networks to avoid binding ambiguities. In order to perceive an accurate perception of the world around us, as binding has to happen on several different levels. For example, parts of an object need to be combined to form an integrated object (part binding), information from different modalities needs to be connected (cross-modal binding) and changing states of objects over time need to be represented as one continuous object (temporal binding) (Treisman, 1996). However, in this literature overview we focus on the binding that happens between several features of an integrated object. For instance, when looking at a desk with a blue notebook and a red pair of scissors, how does our perceptual system realize blue and a rectangular shape belong together when the colors blue and red are both simultaneously activated in early visual areas?

This problem is especially prevalent because sometimes the visual system fails at this challenge. Binding can break down when we attribute features of a different object to another object of interest, thereby perceiving incorrect combinations of features. For example, in a display with a red square and a green triangle people might have wrongly seen a red triangle. This is called an illusory conjunction or feature misbinding. Although illusory conjunctions are rarely consciously encountered in daily life, strict experimental conditions make it possible to reliably elicit them (Prinzmetal, 2015). To give an illustration, when participants had to decide whether a target probe was present during brief presentation of colored letters, participants are more likely to make a mistake when a conjunction probe (i.e., combines two features that are both present in the display, but belong to separate objects) is presented compared to a feature probe (i.e. combines one feature present in the display with another feature that is possible in the experiment, but not present in the display) (Treisman, 1982; Prinzmetal et al., 1986). When participants wrongfully decide the conjunction probe is present in the display, an illusory conjunction has occurred.

Particular situations can sometimes also trigger these illusory conjunctions. First of all, items with adjacent locations to targets will be more likely to exhibit these misbindings compared to items that are spaced further apart (Chastain, 1982; Lasage & Hecht, 1991; Snyder, 1972). Secondly, misbinding of stimuli with physically similar characteristics tend to occur more frequently. For example, when the colors of display letters were more similar, participants were more likely to incorrectly bind these colors to an unrelated display letter (Ivry & Prinzmetal, 1991). Lastly, like is proposed by FIT, attention plays an important role in preventing illusory conjunctions. For example, Prinzmetal et al. (1986) used location cues (with 75% accuracy) to indicate at which location around a point of fixation the target will appear for a probe-recognition task. When the cue was non-informative and thus attention was not directly focussed on the targets, illusory conjunctions were impacted to a higher degree by the lack of attention compared to feature errors.

FIT by Treisman argues for the importance of spatial information in the binding process of integrated objects. Once participants deploy their attention to a specific location, individual features of the object can be bound. From this viewpoint, spatial-attention is the driving force of binding (Robertson, 2003). Bound representations get integrated into an ‘object-file’ that gets merged into a mastermap that contains the spatial and temporal properties of objects (Kahneman & Treisman, 1992; van Dam & Hommel, 2010). Therefore, perception of integrated objects should be mediated through accessing their location. Consistent with this claim, it is demonstrated that directing attention towards the location of target stimuli before these appeared enhanced the perception of their properties (Prinzmetal et al., 1986). In addition, findings show that disruption of spatial attention can lead to problems with detection of objects consisting of multiple features, while this was not found for single-feature-objects (Robertson, 2003; Treisman, 1996). These studies further demonstrate that the ability to guide attention towards the target location is essential for an accurate perception of complex objects.

However, the possibility of location-independent attention to objects has been proposed by several other studies. Interest in location-independent object-based attention started with findings of same-object advantages, as shown in Duncan (1984). This researcher showed participants two overlapping objects, a tilted line on top of a box. Participants were asked to report either one or two features of either the same or different objects. As far as the relevant features, the box stimuli could either differ in size or in the side at which a gap was located. For the tilted line, participants could indicate either the direction of the tiled or the texture of the line (dotted or dashed). Results indicated better feature identification when the features belonged to the same object (e.g., participants reported size of box and side of gap) compared to when they belong to a different object (e.g., participants reported size of box and direction of the line). From this, a system separate from spatial-attention was presumed, an object-based attentional system.

Same-object benefits can also be found in stimulus detection tasks. In an experiment by Egly et al. (1994) participants had to indicate the presence of a target on one of the two bar stimuli presented on either side of the fixation cross. Participants received a pre-cue indicating where the target would most likely appear (with a reliability of 75%). Cues emphasized the outline of half of one of the two bar stimuli. In this way, three possibilities for target positions are created. Targets could either appear at an validly cued location (‘valid target’), appear at an uncued location still inside the boundaries of the cued bar stimulus (‘invalid same object target’), or appear at an uncued location that lays outside the boundaries of the cued bar stimulus (‘invalid different object target’). Crucially, the distance between the ‘valid target’ and both the ‘invalid same object target’ and the ‘invalid different object target’ were equal. Hence, the spill-over of spatial attention caused by an invalid cue should be equal for both targets. As is expected by an object-based attentional system, same-object targets were reported more quickly compared to different object targets. This argues for an attentional system in which attention spreads inside of object boundaries and in which the observation of one feature implies selective attention towards the entire object. This same-object advantage could even be found when participants were not aware of the object boundaries themselves because of masking (Norman et al., 2013).

Furthermore, activation spreads through objects when participants focus on a particular object, even if the object boundaries are outside the focus of spatial attention. This has been demonstrated in flanker tasks where incompatible distractors had a larger effect on target discrimination when both belonged to the same object (Chen & Cave, 2006; Richard et al., 2008). Evidence for an object-based attentional system in perception has additionally been found in spatially superimposed stimuli. Here, attention deployment to one of the superimposed objects enhanced object detection. This can be seen in increased sensitivity (Duncan, 1984; Lui et al. 2007) and after-effects of these stimuli (Boynton et al., 2006; Lankheet & Verstraten, 1995; Melcher et al., 2005). Because spatial attention is shared across these objects, it is concluded that a different kind of attention, object-based attention, also plays an important role in how our attentional mechanisms divide their resources. Taken together, this evidence demonstrates that attentional systems might not only select through spatial selection, but additionally directly select integrated objects for further processing.

Interactions Between Selection by Object and Space

Although space- and object-based attention have been studied extensively, they often are investigated separately from each other. Further research tried to examine how these two attentional systems are related to each other. There are currently two proposed possibilities (Chen, 2012; Soto & Blanco, 2004). Firstly, object-based attention can work in a location-independent way. In this perspective, object features are selected without additionally activating their specific locations. Therefore, space- and object-based attention work in an independent manner and can be applied in accordance with task demands. Secondly, object-based attention could work in interaction with space-based attention. In this case, object features might be selected through mediation of object location. This interaction is caused by the fact that they both rely on the same higher-order mechanism.

Soto and Blanco (2004) studied the interaction between space-based and object-based attention. In contrast to the object-based cues used to indicate bar stimuli in Egly et al. (1994), their experimental design additionally allowed for the target to either appear at a cued or uncued location, separately from object cues. In this way, the influence of space-based and object-based attention can be evaluated separately. Four colored circles were shown to participants, from which one was pre-cued. This cue could either indicate the probable location or the probable object in which a target would appear. After the target appeared, participants had to indicate the orientation of the target object. Crucially, before targets appeared, the colored circles could switch position with their neighbouring circle. Therefore, targets could be at a cued or uncued location, while at the same time being at either a cued or uncued object. Interestingly, they found evidence for both spatial- and object-cueing effects. Moreover, even when the cue was task-irrelevant for space or object, cueing effects could still be found. So, even if participants were instructed that cues would reliably signal the location of targets, facilitation effects could still be found when targets appeared in the same object but different location as the cue. Although, space-based effects seemed stronger compared to object-based effects in this case. Furthermore, object-based effects were only observed when spatial cues turned out to be invalid. These observed interactions between space- and object-based attention seems to indicate an integrated attentional mechanism.

The conclusions of these researchers stands in contrast with others who argue for separate and independent spatial- and object-based attentional mechanisms based on the demands of target representations (Vecera & Farah, 1994; White et al., 2015). In such a view, space- and object-based attention work in parallel ways. For example, White et al. (2015) concluded that color-cues and space-cues had an additive effect when participants had to indicate the visual field in which a saturation increase of overlapping dots appeared. Only when enough competition was added to this task did they find interaction effects. They argued that therefore the underlying mechanisms act independently from one another during normal scene perception.

However, Liang and Scolari (2020) recently investigated the difference between accuracy patterns of valid and invalid single location/color-cues and compared this to combined location + color-cues. Their findings illustrate that when space- and object-based attention are both used for target detection. They each seem to influence different components of decision making and target representation. Thus, there seems to be some higher-order mechanism that makes this interaction possible. Spatial- and object-based attention might also differ from each other in time course, with object-based attention being implemented only after space-based attention (Lui et al., 2007).

To summarize, using pre-cueing experiments two important attentional systems that help us select relevant items in displays can be demonstrated. Like is proposed by FIT, space-based selection plays an important role in target processing. Additionally, object-based attention can be deployed on top of space-based attention. It is proposed that space- and object-based attention interact through a common mechanism and get deployed according to task-demands.

Binding in VWM

When objects are encountered during scene perception, we do not only need to perceive them as one integrated object. Often, it is our goal to actively maintain these object representations even when scene presentation ends. In experimental settings, this need is often implied by using recalling experiments. Moreover, in everyday life object maintenance is also important in creating the perception of continuous scenes. Overall, both our selective visual attention (external attention) and attention involved in VWM processes (internal attention) are of importance in the successful representation of integrated objects. In a first section, the implications of holding objects in VWM will be discussed. A second section will address the common mechanisms between the perception of objects guided by external attention and the deployment of internal attention in VWM maintenance.

The most influential theory of working memory is proposed by Baddley and Hitch (1974), who introduced a working memory model organized according to multiple independent components, called the slave systems. The first slave system, the phonological loop, is involved in the retention of verbal information, while the second slave system, the visuo-spatial sketchpad, is involved in the retention of visual and spatial information. These independent stores are controlled by a general central executive, which supervises and guides these memory stores when needed, like when limited attentional resources need to be allocated to different components. Evidence for such modality-specific and independent stores can be seen in dual-tasks (Baddeley, 1998; Baddeley et al., 2009; Nijboer et al., 2016). For example, early behavioural experiments showed that spatial tasks, like visually tracking a target, interferes with other visual or spatial tasks, but not with tasks dependent on the phonological loop (Baddeley, 1998). So, according to this theoretical model, VWM consists of the information maintained in the visuo-spatial sketchpad. Luck and Vogel (2013) defined some further criteria to define VWM: (1) not only the information, but also the representation itself must be visual in nature, (2) memory content must be actively maintained, this separates it from both iconic (or sensory) memory as well as long term memory, (3) the visual memoranda should be applied during a broad range of cognitive tasks. So, VWM serves to actively hold information intended for manipulation during cognitive tasks.

Limited capacity of visual working memory is another important characteristic of VWM storage. Researchers have long been interested in the question of how much information one can retain over a short period of time. Traditional studies in this domain often look at memory slots as an estimation of the VWM capacity. A slot represents the ability to actively maintain one visual item. These studies made use of the change detection paradigm, in which participants are shown a display that consists of a set of items. Participants are instructed to keep this display in memory. Later, participants get offered an item and have to judge whether or not this was visible in the previous display. Using this paradigm, the capacity of VWM is most often estimated at three to four items (Dai et al., 2019; Luck & Vogel, 1997; Vogel et al., 2001).

This converges with other evidence, such as that provided by ERP experiments. Contralateral Delay Activity (CDA) is a negative slope observed during ERP recordings on posterior sites and contralateral to the visual field in which memory items are presented. This is a very interesting way to measure working memory capacity because this amplitude is found to correspond to VWM capacity, individual differences and capacity limits in participants (Carlisle et al., 2011; Luria et al., 2016). For example, Vogel and Machizawa (2004) demonstrated that the mean amplitude from this signal plateaued when three to four items were stored in working memory. Before this limit was reached, each additional item stored in VWM resulted in an according decrease in amplitude. This could also be seen on an individual level, participants with lower VWM capacities reached this plateau more quickly compared to individuals who scored higher on the VWM tasks. Altogether, this pattern is thought to reflect the capacity occupied during maintenance in working memory. Findings from experiments that measure CDA activity confirm again that VWM capacity lies around three to four items.

Interestingly, we can see this same capacity estimate when one object contains multiple features. Vogel et al. (2001) were the first to compare the VWM capacity for objects with single features to the VWM capacity for objects with a combination of features. Using a change detection paradigm participant’s VWM capacity was again estimated to be three to four items when remembering displays in which either colored squares or oriented bars were presented. First of all, this was taken as evidence that VWM possesses a general capacity limit, rather than having a separate limit for specific feature types. Additionally, researchers presented participants with displays comprised of colored, oriented bars in a multi-feature object condition. These muti-feature objects were found to be memorized just as well as single-feature object, meaning that these muti-feature objects also showed a VWM capacity of three to four items. This ability to combine multiple sources of information into a single memory item, thereby increasing VWM capacity, has been called the binding benefit in VWM.

These studies reveal that VWM representations include integrated objects, not independent stores for each of their features. This is a robust finding and has been replicated for many different types of feature combinations, including color-location conjunctions (Wheeler & Treisman, 2002), color-shape conjunctions (Allen et al., 2006; Xu 2002) and texture-shape conjunctions (Delvenne & Raymond, 2004). Furthermore, experiments using CDA also confirm this finding. Luria and Vogel (2010) showed that the CDA for objects with color-polygon conjunctions and color-orientation conjunctions were less negative compared to when color and polygon type or orientation were represented as separate objects. In short, several lines of research suggest that maintenance in VWM happens based on integrated objects, in which multi-feature objects are stored as efficiently as single-feature objects.

Guiding Attention in Perception and VWM

There is a growing body of evidence which suggests that attention directed towards objects maintained in VWM is similar to selective attention during perception (Awh et al., 2006; Chun, 2011; Chun et al., 2011; Gazzaley & Nobre, 2012; Postle, 2006). From this viewpoint, one central attentional mechanism can be focussed on external stimuli, just as it can be deployed in focussing on the internal representations of these stimuli.

This link between cognitive and perceptual representations can be demonstrated using fMRI decoding studies. These studies show that VWM items can be retrieved when looking at sensory areas (Nelissen et al., 2013; Riggall & Postle, 2012). Similarly, neuroimaging research shows overlapping activity between orienting attention to external locations and orienting attention to internal representations (Nobre et al. 2004; Ruff et al., 2007). These results support the sensory recruitment hypothesis (Lee & Baker, 2016; Postle, 2006), which proposes that regions responsible for the encoding of stimuli are also engaged when representations of these stimuli are activated in working memory. These studies show the intrinsic relationship between the perceptual system and VWM. Likewise, attention also seems to guide what information gets encoded into VWM representations. Both pre- and retro-cues have the ability to facilitate recognition of memory items in change detection tasks (Griffin & Nobre, 2003; Matsukura & Vecera, 2009). Thus, items can be retrospectively selected due to a shift in internal attention. In addition, behavioural experiments show that items held in VWM can influence perceptual attention. For example, Kiyonaga and Egner (2014) showed that storing a color-word in working memory reproduces classic Stroop effect findings when participants performed a color-discrimination task. Other research also indicates that storing an item in VWM guides covert attention, leading to facilitated recognition of similar items in displays (Downing, 2000; Foester et al., 2018; Pashler & Shiu, 1999; Soto & Humphrey, 2009). So, attention seems to play an important role in the encoding and maintenance of VWM representations.

The object-based effects typically found in perceptual studies (Desmone & Duncan, 1995; Duncan, 1984) have also been demonstrated during the retention of integrated objects in working memory. More specifically, Sahan et al. (2019) used fMRI encoding models to retrieve the neural representations of two items held in VWM. Internal object-based attention was manipulated by presenting participants with a relevance cue before they had to reproduce a specific feature dimension (color or motion) of one of the two items on a continuous scale. This relevance cue could indicate which object was relevant for reproduction (bound condition), in which case both of the object dimension of this particular object could be probed. Alternatively, the cue indicates which stimulus dimension was relevant (unbound condition), in which case either the first or the second object could be probed. Results showed that the neural representation of the a non-probed feature dimension were higher in the bound condition compared to the non-bound condition. So, when two features are held in memory, the prioritization of either one of them causes less weakening of the non-prioritized feature when they are bound into an integrated object in VWM. Therefore, binding in VWM provides same-object benefits.

Even more interesting, when VWM is occupied, this directly influences object-based attention effects. Barens et al. (2001) asked participants to judge superimposed stimuli, similar to those of Duncan (1984), during the retention interval of a memory task. Only when the memory task involved object-based memory did the authors find a reduction in the same-object benefit. However, when spatial- or verbal working memory was required, normal same-objects benefits could again be found. Likewise, when selective attention only required spatial attention, such as in a visual search task, object-based attention is not affected (Matsukura & Vecera, 2009). This illustrates the importance of object-based attention for the representation of integrated objects in VWM. These dissociable inference effects of spatial- and object-based VWM interference are consistent with the interactions observed between space- and object-based attention during perception.

In sum, VWM and perceptual attention seem connected through common attentional mechanisms. In this way, attention can be deployed to external stimuli and internal representations in the same way. It is proposed that the same network activated during perception is also necessary to maintain items in VWM (Jaswel, 2012; Nelissen et al., 2013; Riggall & Postle, 2012). Furthermore, the distinction between space- and object-based attention found in visual selection studies also seems to be related to space- and object-based attention in VWM.

Object-Binding or Binding Through Space in VWM

Like is discussed before, to understand perception and processing of integrated objects, both spatial- and object-based attention are important to take into consideration. Likewise, this is an important distinction we can also find in retaining objects in VWM (Matsukura & Vecera, 2009). Substantial research is interested in how integrating multiple features into bound objects is influenced by these attentional processes. A first question to address is whether binding in VWM is an active or passive process. That is, does the maintenance of bound features require more attention compared to the maintenance of constituent features? Several studies show that occupying domain-general attentional processes with a secondary task, such as with backward-counting tasks (Allen et al., 2006; Allen et al., 2012; Brown & Brockmole, 2010) and tone categorization tasks (Morey & Bieler, 2013; Stevanovski & Jolicœur, 2011; Vergauwe et al., 2014) do not interfere with the maintenance of bound objects more compared to the maintenance of single features.

On the other hand, adding secondary tasks that specifically engage object-based attention during retention of bound objects often results in poorer retrieval of objects in memory tasks. Mental rotation (Hyun & Luck, 2007), tracking the motion of transparent superimposed dots (Valdes-Sosa et al., 2000) and feature reporting of superimposed stimuli (Duncan, 1984) are all tasks shown to require object-based attention. Shen et al. (2015) used these tasks to consume participant’s object-based attention during the maintenance phase of a probe-recognition task. Bound color-shape conjunctions, color-direction conjunctions and color-locations conjunctions resulted in worse memory performance under secondary task load. Moreover, this decline in performance was significantly larger compared to the recognition of single features under secondary task load. These differences between bound and simple features could not be found when the secondary task required visual search. They concluded that object-based attention was necessary to rehearse integrated object representations in VWM. Later, these findings have been frequently confirmed by similar studies (Gao et al., 2017; He et al., 2020; Wan et al., 2020). This conclusion again confirms the parallels between internal and external attention processes, where object-based attention is deployed when perceiving integrated objects (Boynton et al., 2006; Egly et al., 1994; Lankheet & Verstraten, 1995; Lui et al. 2007; Mechler et al., 2005), the disruption of this object-based attention diminished the representations of these integrated objects in VWM.

Similarly to the role of location in object-based selection (see section: *Interactions between selection by object and space*), the influence of spatial attention during maintenance of integrated objects is still unclear. To answer this question, Vecera and Farah (1994) used Duncan's (1984) stimuli and presented two multidimensional items to participants either superimposed or spatially separate on either side of a fixation cross. Object-effects (i.e., the advantage of two features belonging to the same object, compared to two features belonging to a different object) were measured for both of the conditions. Detection of two simultaneously presented features was tested. Object-effects were larger when the objects were presented at separate locations compared to superimposed objects. This is consistent with the view that item-location guides object-based attention during perception. Conversely, for a feature identification task, object-effects for separate and superimposed items were equal. Because retrieving object features from VWM was not influenced by the object location, this indicates object-based effects in VWM that are independent from location information. Although, Kramer et al. (1997) argued that location-mediated object selection in VWM can still be possible. They showed that by including a filler stimulus together with the superimposed stimuli during retention, such that both conditions involved two stimuli on either side of the fixation cross, greater object-effects for the separate condition could be found. In this way, spatial-attention is still activated when reporting object features.

The need for spatial attention during maintenance of bound representations in VWM is likewise studied in dual task experiments where the secondary task consumes spatial attention, such as during visual search. While some of these experiments do find an influence of spatial attention interference on the representation of integrated objects (Zokaei et al., 2014), the majority of studies find no additional disadvantage for memorizing objects compared to single features (Johson et al., 2008; Shen et al., 2015).

Whereas dual task studies often differ in attentional manipulation, which causes confusion about the effectiveness of secondary tasks (Fougnie & Marois, 2009), studies using retro-cues can actively direct attention to internal representations of objects. For example, Gajewski and Brockmole (2006) used retro-cues indicating probable recall items of color-shape conjunctions. Recall of object features proceeded by invalid cues were either both reported correctly, or both remembered incorrectly, supporting object-based representations in VWM. Even more interesting however, retro-cues can both indicate target locations or object characteristics. In this way, the internal attention to objects or space can be differentiated more easily (Souza & Oberauer, 2016). Using this paradigm, several experiments concluded that spatial attention to internal representations is not necessary for binding to occur (Allen et al., 2015; Delvenne et al., 2010). However, the question still remains whether object-benefits in VWM are observed because of object-based attention applied to VWM items, or whether these object representations might be mediated by a shared location which is encoded into VWM and in turn enables multiple features to bind to this location.

The Current Study

Several studies have found evidence for object-based representation and binding benefits in VWM (Allen et al., 2006; Delvenne & Raymond, 2004; Luria & Vogel, 2010; Sahan et al., 2019; Xu 2002). The purpose of this experiment is to directly test whether these benefits are due to object-based attention in VWM, or rather are caused by shared spatial attention of these VWM items. If binding of features in VWM is guided by spatial attention, it is predicted that binding only occurs when both of the features share the same space in VWM, irrespectively of whether they are inherently bound at encoding. Alternatively, if features are bound without the need of spatial attention, only stimuli perceived as inherently bound will be stored as integrated items in VWM. To determine this, participants are asked to recall the features of two multiple-feature (color and direction of motion of patches of dots) sample items. The dimensions of these items are always presented in a spatially overlapping way, either because they are inherently bound to the same object (colored moving dots) or because the two features are presented overlapping each other on the same location (static color dots overlapping gray motion dots). The spatial binding hypothesis would predict that the recall precision of both inherently bound and spatially overlapping features is equal, because of their shared spatial characteristics that determine their representation in VWM. As a control, spatial attention in VWM can also be directed to two different spatial locations in VWM (e.g., color of left item and direction of right item). Here, both spatial- and object based attention would not predict binding in VWM. Additionally, we would expect that recall errors will be correlated if binding in VWM items occurs. While previous research was often unable to find these correlations between recall error (Bayes et al., 2011; Fougnie & Alverez, 2011), recently Sone et al. (2021) illustrated these correlations can be found when using simultaneous recall probes instead of separate feature recall. If the spatial binding hypothesis is true, error correlations should occur for both inherently bound features and overlapping features.

Methods

Participants

A total of 51 participants (M_{age} = 18.21, SD_{age} = 0.78, 45 female) were recruited for this experiment. Participants were first bachelor psychology students at Ghent University and received course credits in exchange for participation. This sample size was chosen to ensure sufficient statistical power, while taking into account the number of conditions (four) used in this design and possible data loss due to exclusion of participants. For this purpose, the guideline of Brysbaert and Stevens (2018) were followed, which recommend a minimum of 1.600 observations across participants for every condition in the experiment. Therefore, it was planned to obtain the results of 51 participants, leading to a total of 2.550 observations for every conditions (50 observations for every participant for every condition). Two participants (2 female) were excluded from the final data analysis due to poor recall performance (mean response error < 80° in all conditions). After data exclusion, the current experiment contained 2.450 observation for every condition, therefore ensuring appropriate power to detect existing effects with 49 participants. All participants reported corrected-to-normal color vision and no color vision abnormalities.

Design

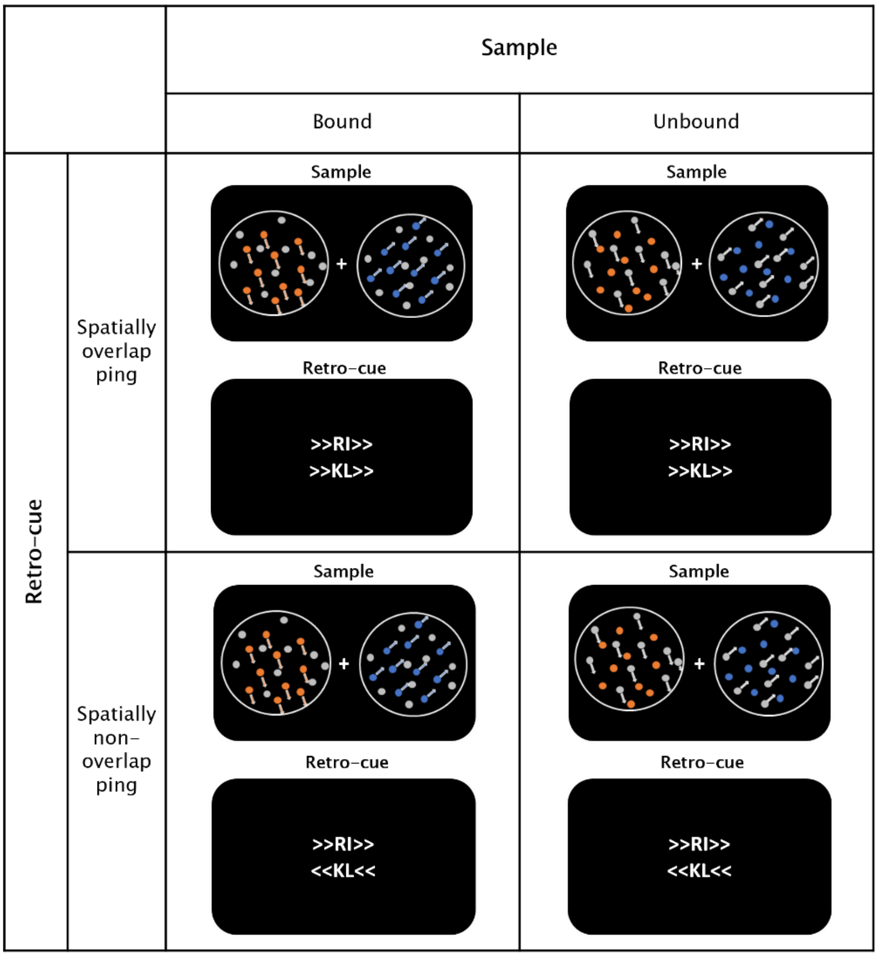
The experiment involved a two-item delayed continuous recall task using retro-cues. At the start of the experiment, participants were simultaneously presented with two sample items. One of these items was presented to the left of the fixation cross and the other one was presented to the right of the fixation cross. Each item consisted of two feature dimensions: ‘color’ and ‘direction of motion’. Crucially, the features belonging to one item could either be presented as ‘bound’ or ‘unbound’. A ‘bound’ item consisted of a patch of dots, where each dot had these same color and direction of motion (i.e., a single dot contained two features). In contrast, an ‘unbound’ item consisted of static colored dots, overlayed by coherently moving gray (uncolored) dots (i.e., a single dot contained one feature). It has been shown that these moving dot stimuli, also known as random dot kinematograms (RDK), are perceived as one coherent object by our visual system (Valdes-Sosa et al. 1998, 2000). Thus, if VWM maintains bound representations of stimuli independently of spatial attention, the features of ‘bound’ stimuli should be integrated into a single VWM item, while the features of ‘unbound’ stimuli are maintained separately. For each sample display, participants were presented with four features (2 colors x 2 directions of motion) which were either unbound or bound into one coherent object.

Participants were instructed to recall two out of the four features held in VWM. Retro-cues were used to select one of the color features (color left/color right) and one of the direction features (direction left/direction right). This was done to avoid confounds of simultaneously holding two items of the same feature dimension in VWM, which has been shown to lead to reduced representations of these same-dimension stimuli compared to different-dimension stimuli (Olsen & Yiang, 2002). The target features were indicated by the presentation of a ‘retro-cue’, informing participants about the dimension type (color/direction) and position (left/right) of both the target features. This retro-cue was added to manipulate spatial-overlap of the features in VWM. An ‘spatially overlapping’ retro-cue selected two features that were presented at the same location (e.g., color left and direction right). In contrast, a ‘spatially non-overlapping’ retro-cue selected two features that were presented at different locations (e.g., color left and direction right).

A 2 (sample type: bound/unbound) x 2 (retro-cue: spatially overlapping/spatially non-overlapping) within-subject design was created. The four conditions of this design are visualized in Figure 1. The presentation of both sample types and retro-cue were balanced across the experiment.

Figure 1

Overview of different conditions according to sample type and spatial overlap



Note. In this figure, movement of the dots is depicted by an arrow indicating the direction of motion. All dots follow the same direction. In a bound sample the colored dots are moving, while the uncolored (gray) dots are static. In an unbound sample the colored dots are static, while the uncolored (gray) dots are moving. Retro-cues depict the dimension (RI = richting (Dutch), direction (English); KL = kleur (Dutch), color (English)) and position (>> = right, << = left) of the cued features.

Stimuli and Experimental Procedure

A trial started with the presentation of a white fixation cross on a black background for 500 milliseconds. Next, the two sample items were presented for 2 seconds. These samples were presented to the left and the right of the centre of the screen. Each sample item contained 400 dots inside of an invisible circle with a diameter of 7.75°. The dots within this sample item were associated with one color and one direction feature. Direction features were randomly sampled between 0° and 360°. Color features were randomly sampled from a 360° CIELAB color space. However, random color and direction values of the second sample item were restricted to values with a minimum of 40° angular separation compared to the randomized features of dots from the first sample item. All dots travelled in the same direction and at a constant speed. Both bound and unbound sample items contained 400 dots. This was done to balance the amount of (feature) information at encoding. For the bound sample, 200 dots contained the randomly selected color and direction feature. Additionally, 200 dots contained a gray value (L = 38, a = 0, b = 0 in CIELAB color space) and remain static. For the unbound condition, 200 dots contained the randomly selected color value, while the remaining 200 overlaying dots contained the randomly selected direction value. In this case, dots with a color value assigned to them remained static, while motion dots contained a grey color value (L = 38, a = 0, b = 0 in CIELAB color space), representing an 'uncolored' stimulus to participants.

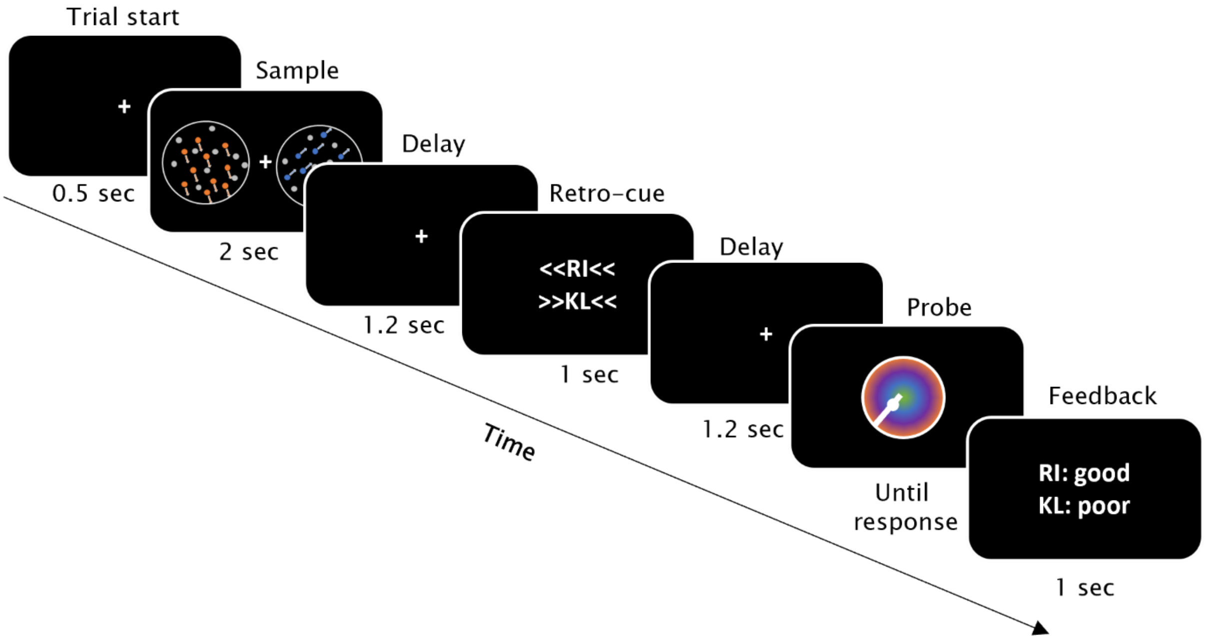
Once sample items have disappeared, a white fixation cross was presented for 1200 ms during a retention interval. Next, a retro-cue was presented on screen for 1 seconds. The retro-cue provided both a dimension indicator and a position indicator to select the cued features. This dimension indicator could either indicate 'Direction' ('Richting' in Dutch, therefore indicated with RI) or 'Color' ('Kleur' in Dutch, therefore indicated with KL), corresponding to the feature dimension that had to be maintained. Furthermore, the position indicator communicated '<< or '>>', corresponding to whether the cued feature dimension needed to be selected from of the left or right sample item. In total, two retro-cues were presented. In other words, participants always had to retrieve two out of the four observed features. Furthermore, the retro-cue indicated the retrieval of one direction feature and one color feature on every trial. Participants were never asked to retrieve two direction features or two color features at the same time. As such, four different dimension-position combinations could be displayed (color left and direction left, color left and direction right, color right and direction right, color right and direction left).

After a delay period of 1200 ms, a probe stimulus was displayed and participants were asked to reproduced the cued features. The nature of these probes is similar to those used in Sone et al. (2021), where bi-feature response probes were used. Participants had to reproduce both features indicated by the retro-cue simultaneously on one probe stimulus. Due to technical limitations, the probes used in this experiment did not require simultaneous indication of the direction and color feature. Rather, direction and color were indicated one after the other, however participants still made use of the same probe stimulus to answer for both features. This probe consisted of a circle (7.75° in diameter) presented in the centre of the screen. The probe had a radial-color gradients using the CIELAB color space. The luminance was kept constant ($L = 80$) while the hue determined the colored rings inside of the circle. Additionally, a white bar (0.05° in width) was presented along the radius of this circle. Participants were able to move the bar 360°, both clockwise and counter clockwise, by moving the computer mouse. Participants were asked to adjust this bar until it represents the direction of motion of the cued item. To confirm their selection of the direction feature participants clicked the right mouse button once. After confirmation, this bar remained on screen in the indicated direction. Once participants had indicated the cued direction, a small white dot (0.1° in diameter) appeared on top of this, now static, white bar. Again, participants could change the position of the dot by moving the computer mouse. The dot could be moved along the length of the white bar (radius of the circle), and moved along the 360 different colors of the CIELAB color space in this way. Participants were asked to adjust the dot to represent the correct cued color value. When participants wanted to register their choice they pressed the right mouse button again. No time pressure was placed on these feature reproductions, however after 30 seconds without an answer the experiment continued. After the feature reproductions, a brief feedback (1 second) was shown to participants. The different feedback options were: ‘great’ (errors ≤ 15°), ‘good’ (errors > 15° and < 30°) or ‘poor’ (errors ≥ 30°). The inter-trial-interval was 1.5 seconds. The interval was indicated by a gray fixation cross. This trial procedure is also illustrated in Figure 2.

During one session, each participant completed 10 blocks of 20 trials (for a total of 200 trials) during a one-hour time slot. During each trial, both color and direction recall were measured. Of these trials, 50 trials measured recall of bound and spatially overlapping features (BO), 50 trials measured recall of bound and spatially non-overlapping features (BNO), 50 trials measured recall of unbound and spatially overlapping features (UBO), and 50 trials measured recall of unbound and spatially non-overlapping features (UBNO). The order in which these trial types appeared was randomized within blocks. Additionally, a block of 20 practice trials was administered before the experiment started.

Figure 2

Example of a trial procedure



Data Analysis

Continuous error values were calculated by comparing the angular distance between the presented feature and the reported feature. All analyses were carried out using RStudio, version 1.2.5033 (RStudio Team, 2020). To analyse the absolute error values a linear mixed effects model (LMEM) was fitted using the lme4 package (version 1.1-23) (Bates et al., 2015) with sample type (bound vs unbound), overlap condition (spatially overlapping vs spatially non-overlapping), and their interaction as fixed effects. The random effects structure of the model was determined by a model building approach. A random slope for sample type, overlap condition and their interaction were iteratively added to the random effects structure of a base model and evaluated to determine their contribution to this model. The base model included only a random intercept for every subject. Therefore, the contribution of a random slope for every predictor was tested separately by comparing the base model to a model including a random slope. This comparison was done by using a Likelihood Ratio Test (LTR) that compared the Akaike information criterion (AIC) of both models. Models which did not perform significantly better than the base model will not be selected. The final model was chosen by selecting the model with the highest AIC compared to the base model. In this way, we are able to avoid overparameterization, like is seen in a maximum random effects structure (Matuschek et al., 2017).

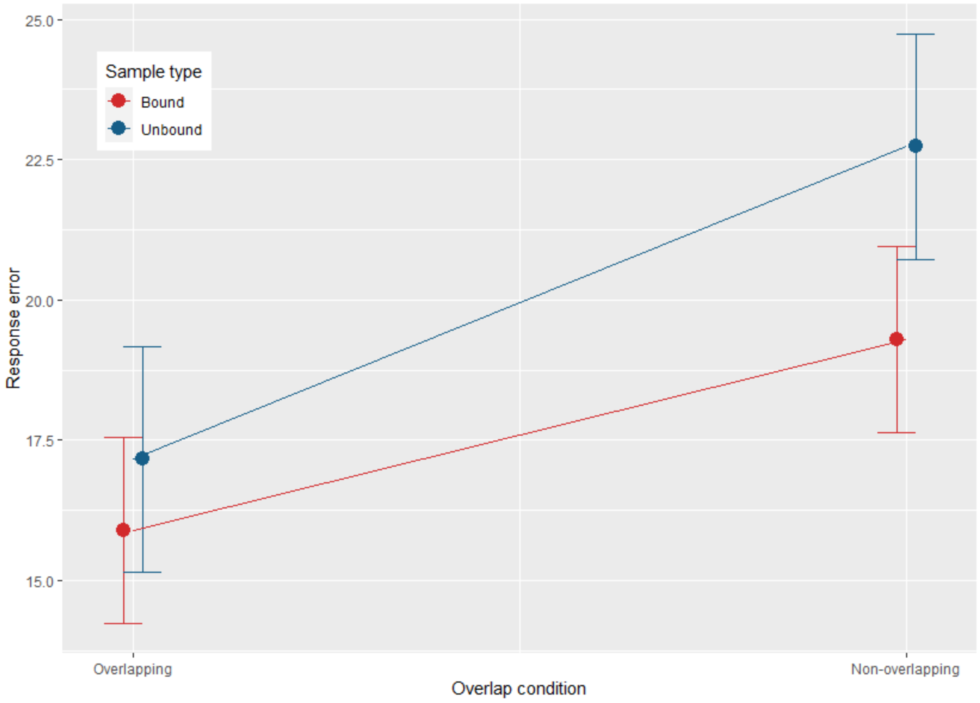
Results

Response Error of Direction Features

Absolute response error for direction recall is shown in Figure 3. Response error was analysed using a linear mixed model with a fixed effect of binding condition (2 levels), overlap condition (2 levels) and their interaction, along with a random effects structure which includes random intercepts and slopes for binding condition for every participant. To analyse the fixed effects of the model a Wald's chi squared test was used. See Table 1 for an overview of the estimates and confidence intervals of the predictors included in the model. The analysis revealed a main effect of overlap condition, $X^2(1, N = 49) = 75.211$, $p < 0.001$. Deviations from the probed direction were lower when both the probed color and direction features were spatially overlapping ($M = 16.525$, $SD = 5.855$) compared to when these features were spatially non-overlapping ($M = 21.011$, $SD = 7.782$). Additionally, a main effect of binding condition was found, $X^2(1, N = 49) = 11.035$, $p < 0.001$. Deviations from the probed direction were lower when sample feature were presented as bound ($M = 17.59$, $SD = 6.111$) compared to when they were presented as unbound ($M = 19.946$, $SD = 8.053$). Furthermore, a significant interaction effect between overlap condition and binding condition was revealed, $X^2(1, N = 49) = 4.376$, $p = 0.036$. Bound samples showed a smaller increase in response error from spatially overlapping to non-spatially overlapping condition (+ 3.403), as compared to unbound samples (+ 5.569). Post-hoc comparisons reveal no significant difference between the response errors in the BO (bound sample, overlapping condition) and UBO (unbound sample, overlapping condition) ($t(48) = -1.645$, $p = 0.107$), while response errors in BNO (bound sample, non-overlapping condition) were significantly smaller in comparison to response errors in UBNO (unbound sample, non-overlapping condition) ($t(48) = -3.287$, $p < 0.01$).

Figure 3

Effect of sample type and overlap condition on direction response errors



Note. Absolute angular distance between the presented direction and the reported direction for each of the four conditions. Error bars represent standard errors.

Table 1

Overview of parameters used in LMEM for analysis of direction response error

		Direction	
Predictors	*Estimates*	*CI*	*p*
(Intercept)	18.77	17.22 – 20.32	**<0.001**
Sample type	1.18	0.48 – 1.87	**0.001**
Overlap condition	2.24	1.74 – 2.75	**<0.001**
Sample type * Overlap condition	0.54	0.03 – 1.05	**0.036**
Random Effects			
σ^2	655.69		
$\tau_{00\ \text{Subj}}$	27.43		
$\tau_{11\ \text{Subj.Sample type}}$	2.88		
$\rho_{01\ \text{Subj}}$	0.46		
ICC	0.04		
N $_{\text{Subj}}$	49		

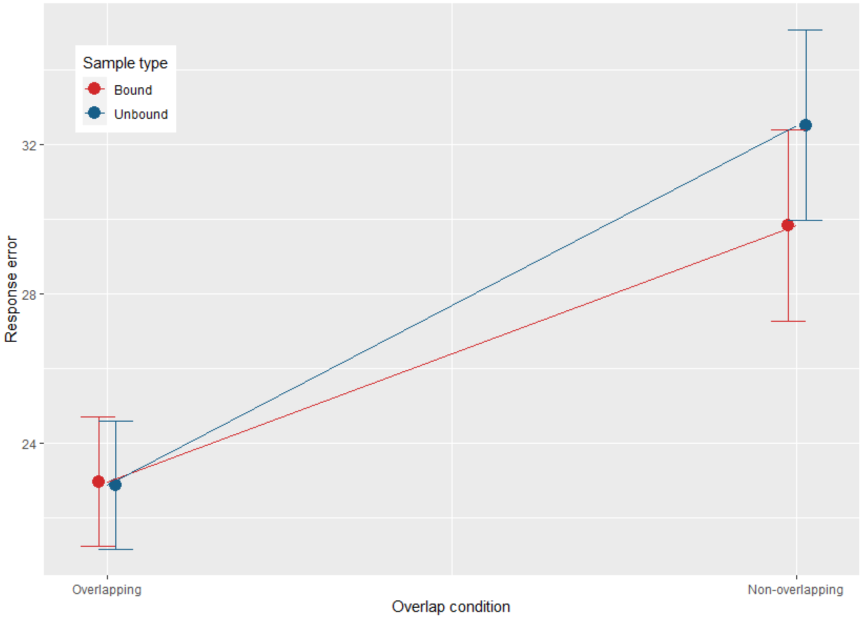
Note. Model equation of LMEM: Direction ~ Sample*Overlap + (1 + Sample | Subj)

Response Errors of Color Features

Absolute response error for color recall is shown in Figure 4. Response error was analysed using a linear mixed model with a fixed effect of binding condition (2 levels), overlap condition (2 levels) and their interaction, along with a random effects structure which included random intercepts and slopes for overlap condition for every participant. To analyse the fixed effects of the model a Wald’s chi squared test was used. See Table 2 for an overview of the estimates and confidence intervals of the predictors included in the model. The analysis revealed a main effect of overlap condition, $X^2(1, N = 49) = 79.307$, $p < 0.001$. Deviations from the probed color were lower when both the probed color and direction features were spatially overlapping ($M = 22.917$, $SD = 5.833$) compared to when these features were spatially non-overlapping ($M = 31.176$, $SD = 9.564$). However, no significant main effect of binding condition was found, $X^2(1, N = 49) = 0.01$, $p = 0.918$. Deviations from the probed color when sample feature were presented as bound ($M = 26.395$, $SD = 8.017$) were similar to deviations when they were presented as unbound ($M = 27.699$, $SD = 9.737$). Lastly, a significant interaction effect between overlap condition and binding condition was revealed, $X^2(1, N = 49) = 5.440$, $p = 0.02$. Bound samples showed a smaller increase in response error from spatially overlapping to non-spatially overlapping condition (+ 6.869), as compared to unbound samples (+ 9.649). Post-hoc comparisons reveal no significant difference between the response errors in the BO (bound sample, overlapping condition) and UBO (unbound sample, overlapping condition) ($t(48)= 0.137$, $p = 0.892$), while response errors in BNO (bound sample, non-overlapping condition) were significantly smaller in comparison to response errors in UBNO (unbound sample, non-overlapping condition) ($t(48) = -2.379$, $p = 0.021$).

Figure 4

Effect of sample type and overlap condition on color response errors



Note. Absolute angular distance between the cued color and the reported color for each of the four conditions. Error bars represent standard errors.

Table 2

Overview of parameters used in LMEM for analysis of direction response error

		Color	
Predictors	*Estimates*	*CI*	*p*
(Intercept)	22.96	21.23 – 24.69	**<0.001**
Sample type	-0.09	-1.74 – 1.56	0.918
Overlap condition	6.87	4.71 – 9.03	**<0.001**
Sample type * Overlap condition	2.78	0.44 – 5.11	**0.020**
Random Effects			
σ^2	868.96		
$\tau_{00\ \text{Subj}}$	20.78		
$\tau_{11\ \text{Subj.Overlap condition}}$	24.78		
$\rho_{01\ \text{Subj}}$	0.45		
ICC	0.05		
N $_{\text{Subj}}$	49		

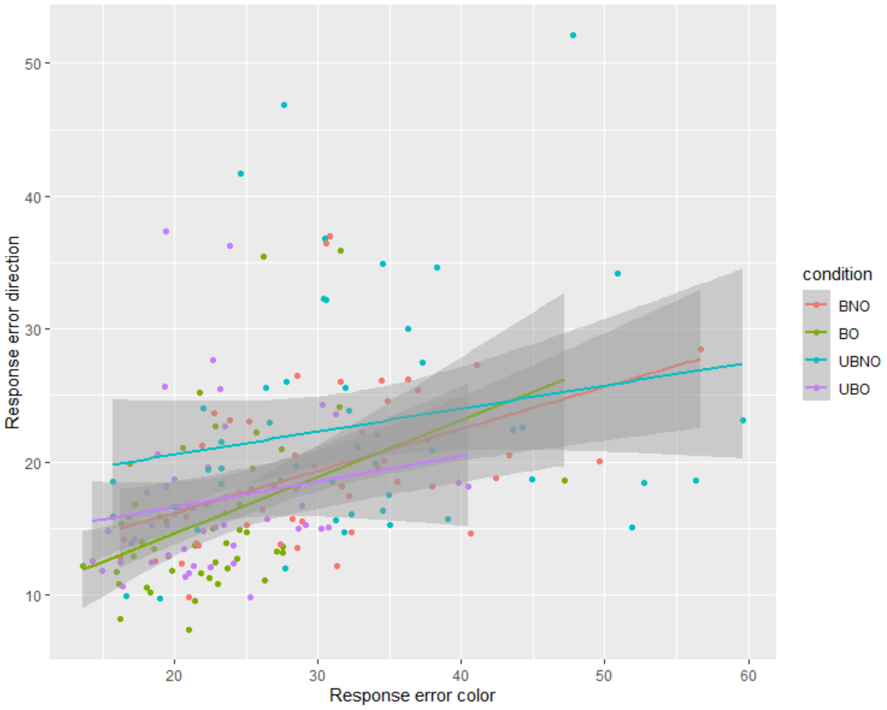
Note. Model equation of LMEM: Color ~ Sample*Overlap + (1 + Overlap | Subj)

Correlation Between Response Errors of Direction and Response Errors of Color

Correlations between color and direction errors for every condition are shown in Figure 5. Overall, a significant correlation of $r(9797) = 0.15$ (95% CI [0.134 – 0.172], $p < 0.001$) was found between response error for direction and response error for color on trial-level. An ANOVA was conducted with fisher's z transformed correlation coefficients as dependent variable. No significant effect of binding condition ($F(1, 48) = 0.011$, $p = 0.919$), overlap condition ($F(1, 48) = 0.639$, $p = 0.428$) or their interaction ($F(1,48) = 0.120$, $p = 0.731$) was found.

Figure 5

Scatter plot and regression line according to condition



Note. Correlation between response errors for color and response errors for direction. BNO = bound sample, non-overlapping condition; BO = bound sample, overlapping condition; UBNO = unbound sample, non-overlapping condition; UBO = unbound sample, overlapping condition. Gray zones represent 95% confidence interval.

Discussion

The current study investigated feature binding in VWM as a consequence of shared location of related features. Additionally, the role of spatial attention in binding was examined by manipulating spatial overlap of the probed features. Our results indicate that features sharing the same location can be indirectly bound by guiding attention to this location in VWM. We found that the response errors of unbound samples features (i.e., color and direction shared one location despite not belonging to the same object) were equivalent to the response errors of bound sample features (i.e., color and direction inherently bound in one object).

This points to an essential role of spatial attention in VWM when binding these features into integrated objects, resulting in a reduced VWM load (Allen et al., 2006; Vogel et al., 2001; Wheeler & Treisman, 2002) and therefore an increase in recall precision (Bays & Husain, 2008; Ma et el., 2014). These results are in line with theories of binding in perception such as FIT (Treisman & Galade, 1980), which state that features are held in separate feature maps. Only when attention is given to a specific location this induces binding of all features present at this location. Furthermore, this would corroborate the general idea that guiding attention in VWM relies on similar principles as the guidance of attention in perception (Chun, 2011; Gazzaley & Nobre, 2012; Postle, 2006). As such, binding in perception and in VWM might rely on similar mechanism as well.

However, the rigid conclusion that spatial attention is strictly necessary for binding in VWM does not fully comply with the results found in the second condition of this experiment, namely the spatially non-overlapping condition. When attention was retro-actively divided across space during recall, we did find an advantage in recall of features from inherently bound samples. Response errors were lower for features presented as inherently bound at encoding compared to when features were presented as unbound during encoding. As such, this seems to suggest that inherently bound stimuli in VWM do differ from features in VWM that were bound through shared location, in such a way that inherently bound features are partly protected from the detrimental effects of dividing attention across different locations in VWM.

When features are bound at encoding, they are held as a single object in VWM (1 + 1 = 1), while the VWM load of features not bound at encoding is higher (1 + 1 = 2).

Competition between these features arises when attention is divided across space. Because the elementary VWM units of bound stimuli are objects (1 vs 1), this decreases the recall errors relative the unbound stimuli (2 features vs 2 features). VWM models of resource allocation can help us to further interpret these differences in response errors. As opposed to slot models, that view the nature of the capacity limit of VWM as a limited amount of spaces (i.e, 3 to 4 locations, slots) to store particular multi-feature items (Dai et al., 2019; Luck & Vogel, 1997; Vogel et al., 2001), resource models propose a division of attention across all stimuli in memory. As a consequence, recall precision decreases for every additional item in memory (Bays & Husain, 2008; Ma et el., 2014).

Thus, when the experimental instructions required division of attention between two locations in the spatially non-overlapping condition, an increase in noise level of the feature representations was seen. This can be observed as increased response error during spatially non-overlapping condition. Consequently, if the precision of features is directly associated with the amount of items focussed on in VWM, we concluded that, while attention is divided across two items during the spatially non-overlapping conditions for both bound and unbound samples, the representations of the unbound sample features seem to be subjected to an even higher level of division of the attentional resources.

As a result, this more divided focus of attention for unbound features exists before retro-cues guide spatial attention to different locations in the spatially non-overlapping condition. In other words, the differences found in VWM representations of bound and unbound samples are already present in the processes occurring before spatial attention is retroactively guided to specific locations, rather than being a function of the processes after allocation of attention.

Automatic Binding as a Benefit of Inherent Object Features

Previous research indicated that binding can result from location sharing of features belonging to the same object (Schneegans & Bays, 2019; Wang et al., 2016). Therefore this spatial-binding hypothesis predicted that the response errors of bound and unbound samples would be alike, both while attention was focussed on a specific location or divided across space. However, from the results obtained in the spatially non-overlapping condition we can conclude that there exists a fundamental difference between the binding of features sharing one location (unbound sample) and the features present in an inherently bound object (bound sample).

Similar to the distinction between inherent binding and binding via location, Eckert and Maybery (2012) investigated the similarity between representations of inherently bound color-shape conjunctions and the binding of shape and background color in VWM. While task-irrelevant features were involuntarily activated during shape- or color-recognition tasks for inherent bindings, this automatic retrieval of task-irrelevant information was not observed when color was merely presented as a background characteristic. These findings are consistent with the distinction between the representation of bound and unbound samples, suggesting a difference in automaticity of binding in VWM based on the inherent binding of features at encoding. While other research identifies proximity and grouping as an important determinant of whether binding in VWM will take place (Woodman & Luck, 2003; Xu, 2006; Xu & Chun, 2007), this does not imply that the mechanisms underlying binding of intrinsic object features and binding features through proximity are identical.

Our results suggest that binding of inherent features occurs independently from attentional resource allocation, as inherently bound samples had an advantage over unbound samples when attention was divided. In the literature, the automaticity of binding of VWM is often investigated by administering a secondary task during the retention interval. In several studies this concurrent attentional load creates deficits in bound representations (Brown & Brockmole, 2010; Fougnie & Marois, 2009; Zokaei et al., 2013). In contrast, the design of this experiment manipulated spatial attention by retroactively dividing attention across space during the non-spatially overlapping condition. For this reason, we specifically target spatial attention, as opposed to more domain-general attentional resources.

While it is revealed that spatial attention is not required for a retro-cue benefit of single-feature objects (Hollingworth & Maxcey-Richard, 2013), it is not yet clear whether spatial attention is necessary for a retro-cueing benefit of the features of multi-feature objects. However, Delvenne et al. (2010) illustrated that the sustained focus of attention elicited by a retro-cue was equally beneficial for individual and integrated features. This observation implies that, before the occurrence of a retro-cue, attention was not guided in VWM to integrate object features. Subsequently, it should be expected that unfocused attention after retro-cues is not detrimental for already established (automatic) feature bindings. This conclusion is in line with the observation that inherently bound samples still outperform unbound samples even under divided attention in this experiment.

Binding Via Shared Location Requires Focused Attention

While the spatially non-overlapping condition revealed a difference between the representation of bound and unbound samples, recall of color and direction during spatially overlapping conditions was equivalent for both sample types. Taken together, these observations suggest that binding of features via shared location can take place only if attention is specifically guided to this location in VWM. Generally speaking, this suggests a separation between a direct, non-attention demanding binding mechanism and an indirect binding mechanism which operates through focussed spatial attention. In accordance with this distinction, Kong and Fougnie (2022) examined the efficiency of retroactively selecting features either through spatial or feature-based cues for VWM updating. Feature-based cues were found to be faster in accessing item representations as opposed to access via location. So, although location cues provided highly accurate information about item representations, feature cues seem to access the memory representations more directly (Li & Saiki, 2015).

Furthermore, this distinction between direct and indirect feature binding can also be found in the different stages of holding items in VWM (Hollingworth & Bahle, 2020; Li & Saiki, 2015; Saiki, 2016). For example, Saiki (2016) used a redundancy gains paradigm, which compares RT of single feature-probes (where either of the two features in a memory display matched the probe) with RT of multi-feature probes (where both features in a memory display matched the probe), while simultaneously using electroencephalogram (EEG) activity to identify different patterns when probes either did or did not share the same location as the objects in the memory display. Results revealed coactivation of both features during VWM maintenance when they belonged to the same object (as opposed to when both features each belong to a different object), regardless of location similarity between the memory display and probe. This indicates that shared location at encoding is essential to form bound representations, while location is not taken into consideration during VWM maintenance. In our experiment, attention was manipulated using retro-cues. As a result of this, spatial attention was not preoccupied during encoding of the samples, allowing for encoding of bound stimuli. Then, only after spatial attention is diverted by use of the retro-cue, location information is made obsolete in order to create direct color-direction bindings. One possible explanation of the results obtained in this experiment is that unbound samples require spatial attention both at encoding and maintenance, thereby only allowing indirect color-direction binding through their shared location when spatial attention stays focussed. In contrast, bound samples may only need spatial attention at encoding, while their representations are location-independent during VWM maintenance.

Study Limitations and Further Research

The amount of features/objects held in VWM in this experiment might impact the generalizability to other multiple-item displays and scenes in daily life. By only holding two (i.e., if each pair of features is bound into one object) or four (i.e., if all features are held separately) items in VWM we barely exceed VWM capacity, which is estimated at three to four items (Dai et al., 2019; Luck & Vogel, 1997; Vogel et al., 2001). Since these items stay below the VWM capacity threshold, there is no need to integrate them into one object, thereby diminishing the strategic advantage of binding in VWM. Although this study shows that features that share the same location do not enjoy the same automaticity of binding as inherently bound features, it is not possible to rule out strategic binding via location to increase VWM capacity. For example, influences of task difficulty were found by Qian et al. (2019), who observed feature-based attentional mechanisms during a low task difficulty and object-based attentional mechanisms during high task difficulty. As such, high task difficulty, implemented as high VWM load, might trigger compensational object-based attentional mechanisms, even when attention is divided.

Additionally, our task instructions did not provide participants with an incentive to bind color and direction features into one object. Rather, participants might be discouraged from doing this due to negative consequences on performance. As such, binding of features could have adverse effects when a retro-cue divides attention in such a way that participants have to report only one of the features from each sample. Importantly, this requires the suppression of the non-cued feature of the same object. It has been illustrated that non-prioritized features are supressed less when belonging to the same object as a prioritized feature compared to when these belong to a different object (Sahan et al., 2019). However, in our experiment the correlation between errors in the reporting of color and direction did not differ for spatially overlapping and spatially non-overlapping conditions. This may indicate that involuntary activation of non-cued features did not play a large role in feature reporting. Still, it would be interesting for further research to compare how the (dis)advantage of binding influences the conditions under which binding occurs, either directly or indirectly via location.

Conclusion

The present study was able to demonstrate equivalent color-direction binding in VWM for bound and unbound samples. This is in agreement with previous research suggesting a location-based binding mechanism in VWM (Schneegans & Bays, 2019; Wang et al., 2016; Zokaei et al., 2014). However, when attention was divided across different locations in VWM we could observe a benefit of bound samples. This indicates a privileged status of bound stimuli in VWM because of their insensitivity to attentional distribution. Taken together, these results suggest a dissociation between two different types of feature binding in VWM. First of all, features can be indirectly bound to each other through their shared location. For this binding to be maintained in VWM spatial attention is necessary. Secondly, features can be directly bound to each other when they form an inherent object at encoding. To maintain this binding in VWM no additional attentional mechanisms are needed.

References

Allen, R. J., Baddeley, A. D., & Hitch, G. J. (2006). Is the binding of visual features in working memory resource-demanding? *Journal of Experimental Psychology: General*, *135*(2), 298–313. https://doi.org/10.1037/0096-3445.135.2.298

Allen, R. J., Havelka, J., Falcon, T., Evans, S., & Darling, S. (2015). Modality specificity and integration in working memory: Insights from visuospatial bootstrapping. *Journal of Experimental Psychology: Learning Memory and Cognition*, *41*(3), 820–830. https://doi.org/10.1037/xlm0000058

Allen, R. J., Hitch, G. J., Mate, J., & Baddeley, A. D. (2012). Feature binding and attention in working memory: A resolution of previous contradictory findings. *Quarterly Journal of Experimental Psychology*, *65*(12), 2369–2383. https://doi.org/10.1080/17470218.2012.687384

Awh, E., Vogel, E. K., & Oh, S. H. (2006). Interactions between attention and working memory. *Neuroscience*, *139*(1), 201–208. https://doi.org/10.1016/j.neuroscience.2005.08.023

Baddeley, A. D., Hitch, G. J., & Allen, R. J. (2009). Working memory and binding in sentence recall. *Journal of Memory and Language*, *61*(3), 438–456. https://doi.org/10.1016/j.jml.2009.05.004

Baddeley, A. (1998). Recent developments in working memory. *Current Opinion in Neurobiology*, *8*(2), 234–238. https://doi.org/10.1016/S0959-4388(98)80145-1

Baddeley, A., & Hitch, G. (1974). The social design of virtual worlds: constructing the user and community through code. *Medical Research Council*, 47–88.

Bahle, B., Thayer, D. D., Mordkoff, J. T., & Hollingworth, A. (2020). The architecture of working memory: Features from multiple remembered objects produce parallel, coactive guidance of attention in visual search. *Journal of Experimental Psychology: General*, *149*(5), 967–983. https://doi.org/10.1037/xge0000694

Barnes, L. L., Nelson, J. K., & Reuter-Lorenz, P. A. (2001). Object-based attention and object working memory: Overlapping processes revealed by selective interference effects in humans. *Progress in Brain Research*, *134*, 471–481. https://doi.org/10.1016/S0079-6123(01)34031-1

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models Usinglme4. Journal of Statistical Software, 67(1). https://doi.org/10.18637/jss.v067.i01

Bays, P. M., & Husain, M. (2008). Dynamic shifts of limited working memory resources in human vision. *Science*, *321*(5890), 851–854. https://doi.org/10.1126/science.1158023

Bays, P. M., Wu, E. Y., & Husain, M. (2011). Storage and binding of object features in visual working memory. *Neuropsychologia*, *49*(6), 1622–1631. https://doi.org/10.1016/j.neuropsychologia.2010.12.023

Boynton, G. M., Ciaramitaro, V. M., & Arman, A. C. (2006). Effects of feature-based attention on the motion aftereffect at remote locations. *Vision Research*, *46*(18), 2968–2976. https://doi.org/10.1016/j.visres.2006.03.003

Brown, L. A., & Brockmole, J. R. (2010). The role of attention in binding visual features in working memory: Evidence from cognitive ageing. *Quarterly Journal of Experimental Psychology*, *63*(10), 2067–2079. https://doi.org/10.1080/17470211003721675

Brysbaert, M., & Stevens, M. (2018). Power Analysis and Effect Size in Mixed Effects Models: A Tutorial. *Journal of Cognition*, *1*(1), 1–20. https://doi.org/10.5334/joc.10

Burwick, T. (2014). The binding problem. *Wiley Interdisciplinary Reviews: Cognitive Science*, *5*(3), 305–315. https://doi.org/10.1002/wcs.1279

Carlisle, N. B., Arita, J. T., Pardo, D., & Woodman, G. F. (2011). Attentional Templates in Visual Working Memory. *The Journal of Neuroscience*, *31*(25), 9315–9322. https://doi.org/10.1523/JNEUROSCI.1097-11.2011

Chastain, G. (1982). Feature mislocalizations and misjudgments of intercharacter distance. *Psychological Research*, *44*(1), 51–65. https://doi.org/10.1007/BF00308555

Chen, Z. (2012). Object-based attention: A tutorial review. *Attention, Perception, & Psychophysics*, *74*(5), 784–802. https://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 and Psychophysics*, *68*(6), 992–1003. https://doi.org/10.3758/BF03193360

Chun, M. M. (2011). Visual working memory as visual attention sustained internally over time. *Neuropsychologia*, *49*(6), 1407–1409. https://doi.org/10.1016/j.neuropsychologia.2011.01.029

Chun, M. M., Golomb, J. D., & Turk-Browne, N. B. (2011). A Taxonomy of external and internal attention. *Annual Review of Psychology*, *62*(February), 73–101. https://doi.org/10.1146/annurev.psych.093008.100427

Dai, M., Li, Y., Gan, S., & Du, F. (2019). The reliability of estimating visual working memory capacity. *Scientific Reports*, *9*(1), 1–8. https://doi.org/10.1038/s41598-019-39044-1

Delvenne, J. F., & Bruyer, R. (2004). Does visual short-term memory store bound features? *Visual Cognition*, *11*(1), 1–27. https://doi.org/10.1080/13506280344000167

Delvenne, J. F., Cleeremans, A., & Laloyaux, C. (2010). Feature Bindings Are Maintained in Visual Short-Term Memory Without Sustained Focused Attention. *Experimental Psychology*, *57*(2), 108–116. https://doi.org/10.1027/1618-3169/a000014

Desimone, R., & Duncan, J. (1995). Neural Mechanisms of Selective Visual Attention. *Annual Review of Neuroscience*, *18*(1), 193–222. https://doi.org/10.1146/annurev.ne.18.030195.001205 de-Wit, L. H., Cole, G. G., Kentridge, R. W., & Milner, A. D. (2011). The parallel representation of the objects selected by attention. *Journal of Vision*, *11*(4), 1–10. https://doi.org/10.1167/11.4.1

Downing, P. E. (2000). Interactions Between Visual Working Memory and Selective Attention. *Psychological Science*, *11*(6), 467–473. https://doi.org/10.1111/1467-9280.00290

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

Duncan, J., & Humphreys, G. W. (1989). Visual Search and Stimulus Similarity Contents. *Perception & Psychophysics*, *54*(6), 716–732.

Ecker, U. K. H., Maybery, M., & Zimmer, H. D. (2013). Binding of intrinsic and extrinsic features in working memory. *Journal of Experimental Psychology: General*, *142*(1), 218–234. https://doi.org/10.1037/a0028732

Eckstein, M. P., Thomas, J. P., Palmer, J., & Shimozaki, S. S. (2000). A signal detection model predicts the effects of set size on visual search accuracy for feature, conjunction, triple conjunction, and disjunction displays. *Perception and Psychophysics*, *62*(3), 425–451. https://doi.org/10.3758/BF03212096

Egeth, H. E., Virzi, R. A., & Garbart, H. (1984). Searching for conjunctively defined targets. *Journal of Experimental Psychology: Human Perception and Performance*, *10*(1), 32–39. https://doi.org/10.1037/0096-1523.10.1.32

Egly, R., Rafal, R., Driver, J., & Starrveveld, Y. (1994). Covert Orienting in the Split Brain Reveals Hemispheric Specialization for Object-Based Attention. *Psychological Science*, *5*(6), 380–382. https://doi.org/10.1111/j.1467-9280.1994.tb00289.x

Foerster, R. M., & Schneider, W. X. (2018). Involuntary top-down control by search-irrelevant features: Visual working memory biases attention in an object-based manner. *Cognition*, *172*(November 2017), 37–45. https://doi.org/10.1016/j.cognition.2017.12.002

Fougnie, D., & Alvarez, G. A. (2011). Object features fail independently in visual working memory: Evidence for a probabilistic feature-store model. *Journal of Vision*, *11*(12), 1–12. https://doi.org/10.1167/11.12.1

Fougnie, D., & Marois, R. (2009). Attentive tracking disrupts feature binding in visual working memory. *Visual Cognition*, *17*(1–2), 48–66. https://doi.org/10.1080/13506280802281337

Gajewski, D. A., & Brockmole, J. R. (2006). Feature bindings endure without attention: Evidence from an explicit recall task. *Psychonomic Bulletin & Review*, *13*(4), 581–587. https://doi.org/10.3758/BF03193966

Gao, Z., Wu, F., Qiu, F., He, K., Yang, Y., & Shen, M. (2017). Bindings in working memory: The role of object-based attention. *Attention, Perception, & Psychophysics*, *79*(2), 533–552. https://doi.org/10.3758/s13414-016-1227-z

Gazzaley, A., & Nobre, A. C. (2012). Top-down modulation: Bridging selective attention and working memory. *Trends in Cognitive Sciences*, *16*(2), 129–135. https://doi.org/10.1016/j.tics.2011.11.014

Griffin, I. C., & Nobre, A. C. (2003). Orienting Attention to Locations in Internal Representations. *Journal of Cognitive Neuroscience*, *15*(8), 1176–1194. https://doi.org/10.1162/089892903322598139

He, K., Li, J., Wu, F., Wan, X., Gao, Z., & Shen, M. (2020). Object-based attention in retaining binding in working memory: Influence of activation states of working memory. *Memory and Cognition*, *48*(6), 957–971. https://doi.org/10.3758/s13421-020-01038-0

Hollingworth, A., & Maxcey-Richard, A. M. (2013). Selective maintenance in visual working memory does not require sustained visual attention. *Journal of Experimental Psychology: Human Perception and Performance*, *39*(4), 1047–1058. https://doi.org/10.1037/a0030238

Hyun, J. S., & Luck, S. J. (2007). Visual working memory as the substrate for mental rotation. *Psychonomic Bulletin and Review*, *14*(1), 154–158. https://doi.org/10.3758/BF03194043

Ivry, R. B., & Prinzmetal, W. (1991). Effect of feature similarity on illusory conjunctions. *Perception & Psychophysics*, *49*(2), 105–116. https://doi.org/10.3758/BF03205032

Jaswal, S. (2012). The importance of being relevant. *Frontiers in Psychology*, *3*(AUG), 1–15. https://doi.org/10.3389/fpsyg.2012.00309

Johnson, J. S., Hollingworth, A., & Luck, S. J. (2008). The Role of Attention in the Maintenance of Feature Bindings in Visual Short-Term Memory. *Journal of Experimental Psychology: Human Perception and Performance, 34*(1), 41–55. https://doi.org/10.1037/0096-1523.34.1.41

Kahneman, D., Treisman, A., & Gibbs, B. J. (1992). The reviewing of object files. *Cognitive Psychology*, *24*(2), 175–219. papers3://publication/uuid/E9F0CEA7-965D-4353-A073-D26F694BDD78

Kiyonaga, A., & Egner, T. (2014). The Working Memory Stroop Effect: When Internal Representations Clash With External Stimuli. *Psychological Science*, *25*(8), 1619–1629. https://doi.org/10.1177/0956797614536739

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*, *126*(1), 3–13. https://doi.org/10.1037/0096-3445.126.1.3

Lankheet, M. J. M., & Verstraten, F. A. J. (1995). Attentional modulation of adaptation to two-component transparent motion. *Vision Research*, *35*(10), 1401–1412. https://doi.org/10.1016/0042-6989(95)98720-T

Lasaga, M. I., & Hecht, H. (1991). Integration of local features as a function of global goodness and spacing. *Perception & Psychophysics*, *49*(3), 201–211. https://doi.org/10.3758/BF03214305

Lee, S. H., & Baker, C. I. (2016). Multi-voxel decoding and the topography of maintained information during visual working memory. *Frontiers in Systems Neuroscience*, *10*(FEB), 1–11. https://doi.org/10.3389/fnsys.2016.00002

Li, Q., & Saiki, J. (2015). Different effects of color-based and location-based selection on visual working memory. *Attention, Perception, and Psychophysics*, *77*(2), 450–463. https://doi.org/10.3758/s13414-014-0775-3

Liang, G., & Scolari, M. (2020). Limited interactions between space- and feature-based attention in visually sparse displays. *Journal of Vision*, *20*(4), 5. https://doi.org/10.1167/jov.20.4.5

Liesefeld, H. R., & Müller, H. J. (2019). A theoretical attempt to revive the serial / parallel-search dichotomy. *Attention, Perception, & Psychophysics*, 228–245.

Liu, T., Stevens, S. T., & Carrasco, M. (2007). Comparing the time course and efficacy of spatial and feature-based attention. *Vision Research*, *47*(1), 108–113. https://doi.org/10.1016/j.visres.2006.09.017

Luck, S. J., & Vogel, E. K. (2013). Visual working memory capacity: From psychophysics and neurobiology to individual differences. *Trends in Cognitive Sciences*, *17*(8), 391–400. https://doi.org/10.1016/j.tics.2013.06.006

Luck, S. J., & Vogel, E. K. (1997). The Capacity of Visual Working Memory for Features and Conjunctions. *Nature*, *390*(1996), 279–281.

Luria, R., Balaban, H., Awh, E., Vogel, E. K., & States, U. (2016). The contralateral delay activity as a neural measure of visual working memory. *Neurosci Biobehav*, 100–108. https://doi.org/10.1016/j.neubiorev.2016.01.003.The

Luria, R., & Vogel, E. K. (2010). Shape and color conjunction stimuli are represented as bound objects in visual working memory. *Neuropsychologia*, *49*(6), 1632–1639. https://doi.org/10.1016/j.neuropsychologia.2010.11.031

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

Matsukura, M., & Vecera, S. P. (2009). Interference between object-based attention and object-based memory. *Psychonomic Bulletin and Review*, *16*(3), 529–536. https://doi.org/10.3758/PBR.16.3.529

Matuschek, H., Kliegl, R., Vasishth, S., Baayen, H., & Bates, D. (2017). Balancing Type I error and power in linear mixed models. *Journal of Memory and Language*, *94*, 305–315. https://doi.org/10.1016/j.jml.2017.01.001

Melcher, D., Papathomas, T. V., & Vidnyánszky, Z. (2005). Implicit attentional selection of bound visual features. *Neuron*, *46*(5), 723–729. https://doi.org/10.1016/j.neuron.2005.04.023

Morey, C. C., & Bieler, M. (2013). Visual short-term memory always requires general attention. *Psychonomic Bulletin and Review*, *20*(1), 163–170. https://doi.org/10.3758/s13423-012-0313-z

Nakayama, K., & Silverman, G. H. (1986). Serial and parallel processing of visual feature conjunctions. *Nature*, *320*(6059), 264–265. https://doi.org/10.1038/320264a0

Nelissen, N., Stokes, M., Nobre, A. C., & Rushworth, M. F. S. (2013). Frontal and parietal cortical interactions with distributed visual representations during selective attention and action selection. *Journal of Neuroscience*, *33*(42), 16443–16458. https://doi.org/10.1523/JNEUROSCI.2625-13.2013

Nijboer, M., Borst, J., van Rijn, H., & Taatgen, N. (2016). Contrasting single and multi-component working-memory systems in dual tasking. *Cognitive Psychology*, *86*(May), 1–26. https://doi.org/10.1016/j.cogpsych.2016.01.003

Nobre, A. C., Coull, J. T., Maquet, P., Frith, C. D., Vandenberghe, R., & Mesulam, M. M. (2004). Orienting attention to locations in perceptual versus mental representations. *Journal of Cognitive Neuroscience*, *16*(3), 363–373. https://doi.org/10.1162/089892904322926700

Norman, L. J., Heywood, C. A., & Kentridge, R. W. (2013). Object-Based Attention Without Awareness. *Psychological Science*, *24*(6), 836–843. https://doi.org/10.1177/0956797612461449

Olson, I. R., & Jiang, Y. (2002). Is visual short-term memory object based? Rejection of the “strong-object” hypothesis. *Perception and Psychophysics*, *64*(7), 1055–1067. https://doi.org/10.3758/BF03194756

Pashler, H., & Shiu, L. P. (1999). Do images involuntarily trigger search? A test of Pillsbury’s hypothesis. *Psychonomic Bulletin and Review*, *6*(3), 445–448. https://doi.org/10.3758/BF03210833

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

Prinzmetal, W. (2015). At the Core of Feature Integration Theory. *From Perception to Consciousness*, *1982*, 211–217. https://doi.org/10.1093/acprof:osobl/9780199734337.003.0018

Prinzmetal, W., Presti, D. E., & Posner, M. I. (1986). Does Attention Affect Visual Feature Integration? *Journal of Experimental Psychology: Human Perception and Performance*, *12*(3), 361–369. https://doi.org/10.1037/0096-1523.12.3.361

Qian, J., Zhang, K., Liu, S., & Lei, Q. (2019). The transition from feature to object: Storage unit in visual working memory depends on task difficulty. *Memory and Cognition*, *47*(8), 1498–1514. https://doi.org/10.3758/s13421-019-00956-y

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

Riggall, A. C., & Postle, B. R. (2012). The relationship between working memory storage and elevated activity as measured with functional magnetic resonance imaging. *Journal of Neuroscience*, *32*(38), 12990–12998. https://doi.org/10.1523/JNEUROSCI.1892-12.2012

Robertson, L. C. (2003). Binding, spatial attention and perceptual awareness. *Nature Reviews Neuroscience*, *4*(2), 93–102. https://doi.org/10.1038/nrn1030

RStudio Team (2020). RStudio: Integrated Development for R. RStudio, PBC, Boston, MA URL http://www.rstudio.com/.

Ruff, C. C., Kristjánsson, Á., & Driver, J. (2007). Readout from iconic memory and selective spatial attention involve similar neural processes. *Psychological Science*, *18*(10), 901–909. https://doi.org/10.1111/j.1467-9280.2007.01998.x

Sahan, M. I., Sheldon, A. D., & Postle, B. R. (2019). The neural consequences of attentional prioritization of internal representations in visual working memory. *Journal of Cognitive Neuroscience*, *32*(5), 917–944. https://doi.org/10.1162/jocn_a_01517

Saiki, J. (2016). Location-Unbound Color-Shape Binding Representations in Visual Working Memory. *Psychological Science*, *27*(2), 178–190. https://doi.org/10.1177/0956797615616797

Schneegans, S., & Bays, P. M. (2019). New perspectives on binding in visual working memory. *British Journal of Psychology*, *110*(2), 207–244. https://doi.org/10.1111/bjop.12345

Shen, M., Huang, X., & Gao, Z. (2015). Object-based attention underlies the rehearsal of feature binding in visual working memory. *Journal of Experimental Psychology: Human Perception and Performance*, *41*(2), 479–493. https://doi.org/10.1037/xhp0000018

Snyder, C. R. (1972). Selection, inspection, and naming in visual search. *Journal of Experimental Psychology*, *92*(3), 428–431. https://doi.org/10.1037/h0032268

Sone, H., Kang, M. S., Li, A. Y., Tsubomi, H., & Fukuda, K. (2021). Simultaneous estimation procedure reveals the object-based, but not space-based, dependence of visual working memory representations. *Cognition*, *209*(December 2020), 104579. https://doi.org/10.1016/j.cognition.2020.104579

Soto, D., & Blanco, M. J. (2004). Spatial attention and object-based attention: A comparison within a single task. *Vision Research*, *44*(1), 69–81. https://doi.org/10.1016/j.visres.2003.08.013

Soto, D., & Humphreys, G. W. (2009). Automatic selection of irrelevant object features through working memory evidence for top-down attentional capture. *Experimental Psychology*, *56*(3), 165–172. https://doi.org/10.1027/1618-3169.56.3.165

Souza, A. S., & Oberauer, K. (2016). In search of the focus of attention in working memory: 13 years of the retro-cue effect. *Attention, Perception, and Psychophysics*, *78*(7), 1839–1860. https://doi.org/10.3758/s13414-016-1108-5

Stevanovski, B., & Jolicœur, P. (2011). Consolidation of multifeature items in visual working memory: Central capacity requirements for visual consolidation. *Attention, Perception, and Psychophysics*, *73*(4), 1108–1119. https://doi.org/10.3758/s13414-011-0099-5

Theeuwes, J. (1996). Parallel search for a conjunction of color and orientation: The effect of spatial proximity. *Acta Psychologica*, *94*(3), 291–307. https://doi.org/10.1016/S0001-6918(96)00003-0

Theeuwes, J., & Kooi, F. L. (1994). Parallel search for a conjunction of contrast polarity and shape. *Vision Research*, *34*(22), 3013–3016. https://doi.org/10.1016/0042-6989(94)90274-7

Treisman, A. (1996). The Binding Problem. *Neurobiology*, 553–565. https://doi.org/10.1002/9781405164535.ch43

Treisman, A. (1982). Perceptual grouping and attention in visual search for features and for objects. *Journal of Experimental Psychology: Human Perception and Performance*, *8*(2), 194–214. https://doi.org/10.1037/0096-1523.8.2.194

Treisman, A. M., & Gelade, G. (1980). A feature-integration theory of attention. *Cognitive Psychology*, *12*(1), 97–136. https://doi.org/10.1016/0010-0285(80)90005-5

Valdes-Sosa, M., Cobo, A., & Pinilla, T. (2000). Attention to object files defined by transparent motion. *Journal of Experimental Psychology: Human Perception and Performance*, *26*(2), 488–505. https://doi.org/10.1037/0096-1523.26.2.488

van Dam, W. O., & Hommel, B. (2010). How object-specific are object files? Evidence for integration by location. *Journal of Experimental Psychology: Human Perception and Performance*, *36*(5), 1184–1192. https://doi.org/10.1037/a0019955

Vecera, S. P., & Farah, M. J. (1994). Does Visual Attention Select Objects or Locations? *Journal of Experimental Psychology: General*, *123*(2), 146–160. https://doi.org/10.1037/0096-3445.123.2.146

Vergauwe, E., Langerock, N., & Barrouillet, P. (2014). Maintaining information in visual working memory: Memory for bindings and memory for features are equally disrupted by increased attentional demands. *Canadian Journal of Experimental Psychology/Revue*

Canadienne de Psychologie Expérimentale, *68*(3), 158–162. https://doi.org/10.1037/cep0000025

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. https://doi.org/10.1037/0096-1523.27.1.92

Vogel, E. K., Woodman, G. F., & Luck, S. J. (2001). Features, Conjunctions, and Objects in Visual Working Memory. In *Journal of Experimental Psychology: Human perception and performance* (Vol. 27, Issue 1, pp. 92–114).

Vogel, E. K., & Machizawa, M. G. (2004). Neural activity predicts individual differences in visual working memory capacity. *Nature*, *428*(April), 1997–2000.

Von Der Malsburg, C. (1999). The what and why of binding: The modeler's perspective. *Neuron*, *24*(1), 95–104. https://doi.org/10.1016/S0896-6273(00)80825-9

Wan, X., Zhou, Y., Wu, F., He, K., Shen, M., & Gao, Z. (2020). The role of attention in retaining the binding of integral features in working memory. *Journal of Vision*, *20*(7), 1–22. https://doi.org/10.1167/JOV.20.7.16

Wang, B., Cao, X., Theeuwes, J., Olivers, C. N. L., & Wang, Z. (2016). Location-based effects underlie feature conjunction benefits in visual working memory. *Journal of Vision*, *16*(11), 1–13. https://doi.org/10.1167/16.11.12

Wheeler, M. E., & Treisman, A. M. (2002). Binding in short-term visual memory. *Journal of Experimental Psychology: General*, *131*(1), 48–64. https://doi.org/10.1037/0096-3445.131.1.48

White, A. L., Rolfs, M., & Carrasco, M. (2015). Stimulus competition mediates the joint effects of spatial and feature-based attention. *Journal of Vision*, *15*(14), 1–21. https://doi.org/10.1167/15.14.7

Wolfe, J. M. (1998). What can 1 million trials tell us about visual search? *Psychological Science*, *9*(1), 33–39. https://doi.org/10.1111/1467-9280.00006

Wolfe, J. M. (2020). Visual Search: How Do We Find What We Are Looking For? *Annual Review of Vision Science*, *6*, 539–562. https://doi.org/10.1146/annurev-vision-091718-015048

Wolfe, J. M., & Horowitz, T. S. (2017). Five factors that guide attention in visual search. *Nature Human Behaviour*, *1*(3), 1–8. https://doi.org/10.1038/s41562-017-0058

Wolfe, J. M., & Horowitz, T. S. (2004). What attributes guide the deployment of visual attention and how do they do it? *Human Perception: Institutional Performance and Reform in Australia*, *5*(June), 73–79. https://doi.org/10.4324/9781351156288-12

Wolfe, J. M., & Bennett, S. C. (1997). Preattentive Object Files : Shapeless Bundles of Basic Features. *Vision Research*, *37*(1), 25–43. https://doi.org/10.1016/S0042-6989(96)00111-3

Wolfe, J. M., Cave, K. R., & Franzel, S. L. (1989). Guided Search : An Alternative to the Feature Integration Model for Visual Search. *Journal of Experimental Psychology: Human Perception and Performance*, *15*(3), 419–433.

Woodman, G. F., & Luck, S. J. (2003). Serial Deployment of Attention During Visual Search. *Journal of Experimental Psychology: Human Perception and Performance*, *29*(1), 121–138. https://doi.org/10.1037/0096-1523.29.1.121

Xu, Y. (2002). Encoding color and shape from different parts of an object in visual short-term memory. *Perception and Psychophysics*, *64*(8), 1260–1280. https://doi.org/10.3758/BF03194770

Xu, Y. (2006). Understanding the Object Benefit in Visual Short Term Memory. *Perception & Psychophysics*, *68*(5), 815–828.

Xu, Y., & Chun, M. M. (2007). Visual grouping in human parietal cortex. *Proceedings of the National Academy of Sciences of the United States of America*, *104*(47), 18766–18771. https://doi.org/10.1073/pnas.0705618104

Zelinsky, G. J. (2008). A Theory of Eye Movements during Target Acquisition. *Psychological Review*, *115*(4), 787–835. https://doi.org/10.1037/a0013118

Zokaei, N., Heider, M., & Husain, M. (2014). Attention is required for maintenance of feature binding in visual working memory. *Quarterly Journal of Experimental Psychology*, *67*(6), 1191–1213. https://doi.org/10.1080/17470218.2013.852232