The human eye has a small region of high acuity at the center of vision, the fovea, that supports fine-grained perceptual processing of objects. We make frequent saccadic eye movements to orient this region to objects of interest in the world. Saccades introduce two problems for the visual system. The first is one of control. How are goal-relevant objects selected as the targets of eye movements over other objects in a scene? The second is one of continuity. Each saccade generates a brief perceptual disruption as the eyes rotate, and the visual information presented on the retina is displaced spatially. How does the eye movement system establish the correspondence between objects visible before and after the saccade to generate the experience of perceptual continuity?

Recent research indicates that two forms of working memory play a central role in solving these problems: visual working memory (VWM) and spatial working memory (SWM), with the combined system termed visuospatial working memory (VSWM). VWM is a limited-capacity system for the active representation of the visual appearance of relevant objects (Luck & Vogel, 2013; Ma, Husain, & Bays, 2014). SWM is a limited-capacity system for the active representation of the locations of relevant objects (Awh & Jonides, 2001). Before a saccade, selection of the saccade goal is strongly guided by the current content of VWM, and a mandatory shift of spatial attention to the saccade target leads to the automatic encoding of the specific features of that object into VSWM. During the saccade, these representations are used to bridge the perceptual gap created by the saccade. After the saccade, VSWM is used to establish object and location continuity and to correct possible errors in saccade landing position. Thus, we argue that VSWM should be conceptualized as part of a closely integrated system for orienting gaze.

VWM and Oculomotor Control

To behave adaptively, we must direct our gaze in a goal-driven manner. For example, when cooking, one must fixate each of the ingredients and utensils as they become relevant (Land & Hayhoe, 2001). Recent developments indicate a central role for VWM in this type of orienting. 

Abstract

Humans make frequent movements of the eyes (saccades) to explore the visual environment. Here, we argue that visuospatial working memory (VSWM) is a fundamental component of the eye movement system. Memory representations in VSWM are functionally integrated at all stages of orienting: (a) selection of the target; (b) maintenance of visual features across the saccade; (c) the computation of object correspondence after the saccade, supporting the experience of perceptual continuity; and (d) the correction of gaze when the eyes fail to land on the intended object. VSWM is finely tuned to meet the challenges of active vision.

Keywords

visuospatial working memory, visual working memory, spatial working memory, eye movements, oculomotor selection, visual attention

Stefan Van der Stigchel and Andrew Hollingworth

1Department of Experimental Psychology, Helmholtz Institute, Utrecht University, and 2Department of Psychological & Brain Sciences, The University of Iowa

Visuospatial Working Memory as a Fundamental Component of the Eye Movement System

Corresponding Author:
Stefan Van der Stigchel, Utrecht University, Department of Experimental Psychology, Heidelberglaan 1, Room H.030, 3584 CS Utrecht, The Netherlands
E-mail: s.vanderstigchel@uu.nl
of eye movement control (or oculomotor control), encapsulated in an experiment by Bahle, Matsukura, and Hollingworth (2017; Figs. 1a and 1b). Participants searched for a target object in a scene. Simultaneously, they maintained a secondary color in VWM for a memory test. Saccade target selection was guided by a representation of the search target, with the eyes directed very efficiently to that object. But it was also guided by the incidental content of VWM: A critical distractor in the scene was more likely to be fixated when it matched the secondary color than when it did not.

First, this experiment illustrates that the strategic maintenance of a target representation in VWM introduces strong control over where the eyes are oriented. In more simplified displays, saccade target selection can be limited, almost exclusively, to objects matching VWM content (Beck, Hollingworth, & Luck, 2012). Second, the active maintenance of a representation in VWM is often sufficient to implement guidance; gaze is directed to memory-matching objects that are known to be irrelevant (Soto, Humphreys, & Heinke, 2006). Third, guidance by VWM tends to dominate other forms of guidance, particularly at the early stages of visual search. In Bahle et al., differential fixation of the critical distractor was observed from the very first saccade on the scene and was not influenced by whether or not the distractor appeared in a plausible location for the target, indicating that VWM-based guidance is implemented before guidance based on scene gist recognition (cf. Wolfe, Võ, Evans, & Greene, 2011). Finally, guidance can be implemented by multiple items in VWM simultaneously: Oculomotor selection was influenced both by a VWM representation of the search target and by a VWM representation of the secondary color. The finding of multiple-item guidance (see also Beck et al., 2012) contrasts with recent claims that only one item can be maintained in an active state in VWM and guide selection (Olivers, Peters, Houtkamp, & Roelfsema, 2011).

In sum, what one happens to be representing in VWM has a substantial influence over where gaze is directed. One locus of this interaction appears to be relatively early within the visual processing stream. Hollingworth, Matsukura, and Luck (2013) examined the influence of VWM on the rapid generation of reflexive saccades to single targets. VWM influenced both the latency and the accuracy of saccades generated in less than 150 ms. A plausible mechanistic implementation of such effects was outlined in Schneegans, Spencer, Schöner, Hwang, and Hollingworth (2014). In this model, VWM maintenance involves sustained activation of subpopulations of neurons in sensory cortex. This activity interacts with the first, feed-forward sweep of sensory input to increase the perceptual salience of items matching VWM content (Gayet et al., 2017), thereby biasing the competition between objects for selection.

### Presaccadic Encoding Into VWM

When saccadic competition has been resolved and the eye movement is programmed, the visual system needs to generate a robust visual representation that can survive perceptual disruption and interference from postsaccadic sensory input. Because this transsaccadic representation depends on the VWM system (for an extensive review, see Irwin, 1992b), it tends to be limited to a subset of scene information, and there is a strong bias to select objects at the saccade target location (Currie, McConkie, Carlson-Radvansky, & Irwin, 2000). Specifically, visual attention shifts to the location of the impending saccade target (e.g., Deubel & Schneider, 1996), leading to the preferential encoding of visual information from that region into VSTM, both in terms of the probability of encoding (Irwin, 1992a) and the precision of the target representation relative to other remembered items (Bays & Husain, 2008). In addition to a close relationship between saccades and VWM encoding, saccade preparation also prioritizes the retention of objects already maintained in VWM (Hanning, Jonikaitis, Deubel, & Szinte, 2016; Ohl & Rolfs, 2017), with the selective retention of items that were originally encoded at the saccade target location.

Further, several recent studies have indicated that oculomotor selection effects on VWM encoding and maintenance are automatic and are specifically related to saccade preparation (Ohl & Rolfs, 2017; Schut, Van der Stoep, Postma, & Van der Stigchel, 2017; Shao et al., 2010; Tas, Luck, & Hollingworth, 2016). In Schut et al. (Figs. 1c and 1d), the precision of VWM for shapes was assessed either with or without an intervening eye movement task. The demand to execute a saccade introduced substantial interference with VWM maintenance. The drop in precision was approximately equivalent to the loss of one object’s worth of information, presumably caused by saccade target encoding into VWM. These selection effects are specific to VWM and to the situation in which a saccade must be executed to a visible object. Saccades produced no interference with working memory for verbal stimuli in Schut et al.’s experiment. And in a similar study by Tas et al. (2016), there was no interference with VWM if participants executed a saccade to empty space or if participants were required to shift attention covertly to a peripheral object without executing a saccade. This latter finding indicates an important dissociation between covert and overt attention and is explained naturally by the idea that the functional relationship between attention and VWM encoding is produced by the demands of saccade execution. Covert shifts of attention do not introduce a perceptual disruption or shift in retinal input. It is only when a saccade has been programmed and will be executed that VWM encoding is required to bridge...
In Experiment 1 of Bahle et al. (a), participants saw a target object and a label that cued the relevant object in the upcoming scene. Then they saw a color to be remembered for a memory test at the end of the trial. Participants searched for the target object and reported the orientation of a small letter “F” superimposed on it. Finally, they completed a forced-choice, within-category color memory test. The color of the memory item was manipulated so that it did or did not match the color of a critical distractor in the scene (in this example, a pumpkin). Eye movements were recorded. Scan paths (b) are shown for the 10 participants who saw this scene item in the mismatch condition and the 10 participants who saw this item in the match condition. Lines represent saccades; circles represent fixations. Note that gaze started at the center of the scene and was ultimately directed to the target object. However, the critical distractor was more likely to be fixated when it matched the secondary color in visual working memory (VWM). Across participants and scene items, the mean probability of critical distractor fixation was .18 in the mismatch condition and .40 in the match condition. In Experiment 1 of Schut et al. (c), participants saw one, two, or three common shapes (a single shape is shown here) and remembered the width-height ratio of each (i.e., the extent to which the shape was stretched in the vertical or horizontal dimension). During the retention interval, participants were asked to complete a task on half of the trials (shown here) that required them to execute a saccade to an object that was briefly enlarged. Finally, participants manipulated the width-height ratio of a test item until it matched the corresponding memory item, allowing an estimate of memory precision. Memory precision on this task (d) is shown as a function of the number of memory items and of whether there was or was not an intervening saccade task (error bars show 95% confidence intervals). Note that the saccade task introduced a drop in memory precision approximately equivalent to the addition of one object to the VWM load.
perceptual disruption and maintain object continuity across retinal displacement.

**VWM Across Saccades Supports Perceptual Continuity**

VWM has been central to accounts of transsaccadic continuity that stress a primary role for a representation of the saccade target object (e.g., Currie et al., 2000; Deubel, Schneider, & Bridgeman, 1996; Irwin, McConkie, Carlson-Radvansky, & Currie, 1994). According to this view, transsaccadic object continuity and the experience of stability is accomplished by a saccade-target mapping operation, in which VWM for presaccadic target properties is compared with postsaccadic sensory input near the fovea. This mapping operation occurs in the context of a strong bias to assume that the world has remained stable (Atsma, Maji, Koppen, Irwin, & Medendorp, 2016; Deubel et al., 1996).

What is the nature of the saccade target information supporting these processes? Both the locations and surface feature properties of objects are used to compute target correspondence: Spatial displacement of the saccade target across the eye movement (Bridgeman, Hendry, & Stark, 1975) and significant changes in surface features (Tas, Moore, & Hollingworth, 2012) both interfere with the perception of a single, continuous object. In addition, the mapping operation is inherently predictive, both for location and for surface features.

In the former case, the attended retinotopic locations of a few objects are updated to account for the spatial displacement (Boon, Belopolsky, & Theeuwes, 2016; Rolfs, Jonikaitis, Deubel, & Cavanagh, 2011), although attention may linger briefly at the old retinal location immediately after the saccade (Golomb, Pulido, Albrecht, Chun, & Mazer, 2010). This type of attentional remapping can be considered dependent on the SWM system given evidence indicating a close relationship between spatial attention and SWM (Awh & Jonides, 2001). In the feature domain, Herwig and Schneider (2014) showed that the representation of the saccade target is influenced by the expected appearance of that object when the eyes land. Specifically, they trained participants to associate different spatial frequencies for the saccade target before and after the saccade. This association then came to bias perceptual experience of the target before the saccade toward the expected postsaccadic spatial frequency.

Given that a representation of the saccade target is maintained across the saccade, to what extent is this integrated with perceptual information activated when the eyes land? Several recent studies have indicated some level of perceptual integration. For instance, when colors were shifted imperceptibly during a saccade (Oostwoud Wijdenes, Marshall, & Bays, 2015), participants tended to report a color that was between the pre- and postsaccadic values. Interestingly, the weights given to pre- and postsaccadic representations in the integrated representation appear to be influenced by sensory uncertainty (e.g., as introduced by acuity differences or the level of visual noise), with more weight given to the more reliable representation (Ganmor, Landy, & Simoncelli, 2015; Wolf & Schütz, 2015).

For such integration to play a functional role in the perception of object continuity, it needs to occur immediately after the saccade. To measure the time course of integration, Fabius, Fracasso, and Van der Stigchel (2016) used a motion illusion (the high phi illusion; Wexler, Glennerster, Cavanagh, Ito, & Seno, 2013), in which an annulus with a random texture rotates slowly and is then replaced by annuli with several different textures (transients). With sufficient rotation duration, participants report a transient as a large rotational jump in the opposite direction. In Fabius et al., the texture rotated in the periphery, and participants executed a saccade to it. Participants observed the illusion if the transient was presented as soon as the eyes landed, indicating that the presaccadically acquired information influenced perception immediately after the saccade.

**VWM Supports Selection and Gaze Correction After the Saccade**

Eye movements often fail to land on the saccade goal, and a corrective saccade is needed to orient gaze to the original target. The process of gaze correction serves as a microcosm of the processes discussed so far, illustrated in experiments by Hollingworth, Richard, and Luck (2008). Participants executed a saccade to a target disk in a circular array of disks that differed only by color. On some trials, the array rotated so that the eyes landed between the target and an adjacent distractor disk. This was meant to simulate the common situation in which the eyes miss the saccade target, and there are several candidate objects near the landing position. Memory for the color of the target allowed participants to efficiently correct their gaze, and the effect of color match was stronger than the effect of relative proximity of the landing position to each of the two objects: Feature correspondence dominated spatial correspondence. Moreover, a secondary VWM load impaired gaze correction, and the features for the secondary task interfered with correction when they were associated, postsaccadically, with the distractor (Hollingworth & Luck, 2009), implicating VWM in solving the gaze-correction problem. Thus, after the primary saccade, the VWM representation of the saccade target is used to establish correspondence, and if there
Van der Stigchel, Hollingworth

Active Representations in VWM

Search Target

Fig. 2. Proposed function of visual working memory (VWM) during eye movement orienting (note that we do not depict representations of spatial working memory). (a) During the task of baking, a lemon is needed, and the person forms a VWM representation of a canonical lemon as the search target to guide selection. (b) Spatial attention shifts to an object that is a relatively close match to the search target representation, in this case butter (highlighted by an orange circle). The butter is selected as the target of the next saccade, and the features of that object are encoded into VWM. (c) A saccade is executed but fails to land on the target. The VWM representation of the saccade target is used to establish object correspondence and to localize the target, leading to a rapid corrective saccade. (d) Detailed inspection of the fixated object reveals that it is not the target, and attention shifts to another object that is also a relatively close match to the search target representation (highlighted again by an orange circle). The features of the next saccade target are then encoded into VWM.
is no appropriate object at the fovea, this representation is used as a template to guide an extremely rapid visual search operation, supporting selection of the original saccade target. For such a mechanism to work, the presaccadic attentional shift should be determined by the planned saccade target location and not by the actual saccade landing position. Indeed, this appears to be the case (Deubel & Schneider, 1996). For example, when there is strong competition for oculomotor selection and the eyes generally land between a target and a distractor, spatial attention is nevertheless allocated to the object that is the goal of the saccade (Van der Stigchel & de Vries, 2015). This distribution of attention would support the encoding of target properties for saccades that will not ultimately land on that object.

**Conclusion**

VSWM is functionally integrated with oculomotor mechanisms at all stages of the orienting process (see Fig. 2). Before the saccade, the following occurs:

- The saccade goal is selected, to a significant degree, by the current content of VWM, with attention biased toward memory-matching objects.
- The presaccadic shift of spatial attention to the target leads to the encoding of target features into VWM in a manner that predicts the postsaccadic appearance of that object.
- Attentional pointers in SWM are then predictively updated to the future retinal locations to maintain attention on goal-relevant objects.

During the saccade, the content of VWM is used to bridge the perceptual gap created by the saccade in a format that is resistant to masking from postsaccadic sensory input. After the saccade, the following occurs:

- Surface feature and position information in VWM and SWM are used to establish the correspondence between a few relevant objects that were visible before and after the saccade, particularly for the saccade target.
- The VWM representation of the target can be integrated with new perceptual input to form a composite representation.
- If the eyes fail to land on the target, the VWM representation is used as a search template, guiding a corrective saccade to the target in a manner similar to the original selection of that object.

We argue that many of the basic properties of VSWM can be understood as arising from the optimization of oculomotor control. For example, the relationship between spatial attention and VWM encoding can be understood as reflecting the demand to bridge the perceptual gap introduced by saccades. The close relationship between visual attention and SWM can be understood as arising from a need to maintain attention on relevant locations across saccadic disruption and delay. The guidance of attention by multiple VWM representations can be understood as reflecting the simultaneous demand to select the ultimate target and to establish correspondence and correct gaze for each of the individual objects fixated during search (which may not always be a precise match to the target). In sum, the oculomotor system does not simply exploit a general working memory system; VSWM is finely tuned to meet the demands of active vision.

**Recommended Reading**


Luck, S. J., & Vogel, E. K. (2013). (See References). One view of the capacity limits of visual working memory in terms of slots as discrete units.

Ma, W. J., Husain, M., & Bays, P. M. (2014). (See References). An alternative view on the capacity limits of visual working memory in terms of available resources.


**Declaration of Conflicting Interests**

The author(s) declared that there were no conflicts of interest with respect to the authorship or the publication of this article.

**Funding**

This work was supported by grants from the Netherlands Organization for Scientific Research (Vidi Grant 452-13-008) to S. Van der Stigchel and from the National Institutes of Health (R01EY017356) to A. Hollingworth.

**Notes**

1. The framework presented here does not depend critically on whether VWM capacity is conceptualized as an item limit or as a flexible resource.
2. Neurophysiological evidence consistent with preferential sampling at the saccade target location was reported by Zirnsak, Steinmetz, Noudoost, Xu, and Moore (2014), who observed the
convergence of receptive fields in frontal eye field neurons toward the planned target location.

References


