An embodied account of visual working memory

Stefan Van der Stigchel

Published online: 19 Mar 2020.

Submit your article to this journal ⊡

Article views: 665

View related articles ⊡

View Crossmark data ⊡

Citing articles: 1 View citing articles ⊡

Routledge
Taylor & Francis Group

An embodied account of visual working memory

Stefan Van der Stigchel

Experimental Psychology, Helmholtz Institute, Utrecht University, Utrecht, Netherlands

ABSTRACT
Traditional models of visual memory rely solely on internal memory and ignore our reliance on the information that is physically present in the external visual world. Experiments on visual working memory generally use paradigms that are designed to maximally load internal memory storage, although these situations do not necessarily translate to the actual use of visual working memory in daily life. Here, I discuss an embodied view of visual memory in which there is a continuous decision about which information to internalize and which information to leave in the external world for (possible) access later in time. In this view, the known limited capacity of visual working memory is not a problem in daily life, as the external world typically remains readily available and can be accessed relatively easily by executing eye movements to relevant locations.

Whenever you walk in a forest or stroll through a downtown area, you experience a rich visual world. You enjoy the various shades of green of the trees or are overwhelmed by all of the details that a busy city centre offers. Although you might have the impression of a rich visual world, we now know that your brain only represents very little of this visual world at each individual moment in time. For instance, remarkably large changes in the environment generally go undetected (Rensink et al., 1997), indicating that the external visual world is only partly represented internally.

Current memory models refer to "visuospatial working memory" as the memory system responsible for the internal representation of the visual world (Logie, 1995). Visuospatial working memory is divided into a visual and a spatial component, with spatial working memory maintaining relevant locations in the visual world and visual working memory (VWM) maintaining visual features of objects (Baddeley & Hitch, 1974). VWM is generally regarded as a capacity-limited, effortful storage for visual information that is no longer available (hence the term "memory").

The recent rise of interest in VWM has resulted in lively debates and important findings on the nature of its capacity (Luck & Vogel, 2013; Ma et al., 2014).

One of the most intriguing issues has been the maximum capacity of VWM. This question has resulted in a fierce debate about whether the capacity limits of VWM should be interpreted in terms of slots as discrete units (Luck & Vogel, 2013) or in terms of available resources (Ma et al., 2014).

Although this is an interesting theoretical discussion, the corresponding experiments might not translate to the actual use of VWM in daily life. It is actually quite difficult to come up with a task in daily life that involves holding multiple visual items in memory, besides perhaps complex visual imagery. Simply look at the effort that participants have to deliver to perform our experiments in the laboratory and it becomes clear why we prefer not to maintain multiple items in working memory. First, maintaining an item in VWM is expensive, as this process requires attentional resources (Cowan, 2005). Second, internal representations are fragile, prone to decay or disturbance due to incomplete or incorrect encoding (Baddeley & Hitch, 1974).

Most experiments on VWM study memory performance for visual information that is no longer physically present. For instance, in change blindness experiments, the observer must identify a change from one visual scene to another, when the initial scene is no longer present (Simons & Levin, 1997). In order to

CONTACT Stefan Van der Stigchel ✉ s.vanderstigchel@uu.nl 🖃 Experimental Psychology, Helmholtz Institute, Utrecht University, Utrecht, Netherlands

correctly perform this task, information has to be stored in VWM. Furthermore, the traditional paradigm for studying the neural correlates of VWM is a task in which an array of random items is presented and removed, and recall is tested a few seconds later (e.g., Gayet et al., 2017; Harrison & Tong, 2009). By enforcing a strategy in which observers have to store information in VWM to correctly perform the task, researchers have ignored the fact that outside of the lab, our external visual world typically remains available and is relatively stable.

So, although the maximum capacity of VWM might be considered to be about 3–4 items when interpreting the limits in terms of discrete slots, this capacity might not be used in daily life when we interact with physically present information. In these situations, humans can depend on the external world to access visual information in their environment. There is no need for an internal representation of multiple objects as long as the visual information is readily available in the external world.

The idea of the world as an external memory is not an original viewpoint in vision science. The idea is perhaps best known from Kevin O'Regan's classic paper (1992) in which he based his ideas on a long tradition of scientists who emphasize the importance of considering the world as an external memory (Haber, 1983; MacKay, 1967; Turvey, 1977). Although the idea might not be new, I consider it worth discussing the viewpoint again in the light of the current debates in the VWM literature. Especially the current focus on the maximum capacity of VWM and the use of paradigms in which information is no longer physically present seems at odds with the view of the world as an external memory.

The focus on the maximum capacity in experiments on VWM is similarly present in the neuropsychological tests that are available to assess a person's maximum capacity of VWM (Berch et al., 1998; Kessels et al., 2000). If we hardly ever use the maximum capacity in daily life, this sort of assessment will have little predictive value of a patient's functioning during daily activities requiring VWM, such as navigation and visual search.

So, how are we able to survive with such a small capacity to internally store information from the external visual world? Luckily, we have a system in place that allows us to internalize very little of the external visual world in internal memory: the eye movement system.

We do not passively perceive the world, but interact with our environment. One of these interactions is the execution of eye movements to relevant locations in our visual world. Because these locations contain objects, we can access these objects by moving our eyes to these locations. Eye movements are so efficient that they allow us to use the world as an external memory (Ballard et al., 1995; Hayhoe & Ballard, 2005; Triesch et al., 2003). Eye movements have even been claimed to be "cheap" as they are executed extremely quickly and are associated with low effort (e.g., Theeuwes, 2012). We are generally not aware of the many eye movements we execute, even though previous research has unravelled that the selection of where to execute the next eye movement is far from trivial (Van Zoest et al., 2017). Despite the multitude of processes necessary to execute eye movements, we generally do not come home after a day of work complaining about how tiring it has been to execute the thousands of eye movements during an average day.

This efficiency of eye movements also underlies our subjective impression of a complete internal representation of the visual world. Because eye movements and attention are tightly coupled (Corbetta, 1998; Van der Stigchel & Theeuwes, 2007), saccades will generally be accompanied by a preceding covert shift of attention to the fixated location. The impression of consciously seeing everything is due to these shifts of attention. Whenever you want to scrutinize an object in your visual world, the resulting saccade will be accompanied by a mandatory shift of visual attention to this object to allow further inspection (Deubel & Schneider, 1996; Van der Stigchel & Theeuwes, 2005). Everything you check on will therefore be immediately available for scrutiny through a quick shift of visual attention, creating the impression of a complete representation of the visual world. The only requirement for such a system to be efficient is an internal memory of where important information is positioned. The use of the world as an external visual memory is therefore enabled by the human ability to make rapid eye movements to relevant locations and to maintain the locations of important information in internal spatial working memory.

Our brain can internalize information from the external visual world by directing the eyes and transferring the information at the fixated location to VWM ("sampling information"), consistent with the

idea that working memory is an emergent property and is strongly associated with different cognitive and motor functions, such as the execution of eye movements (Postle, 2006). We recently concluded that the overlap between VWM and the eye movement system is even stronger than previously thought (Van der Stigchel & Hollingworth, 2018): every time we make a goal-directed saccade, the saccade target is automatically transferred into VWM (Hollingworth & Luck, 2009; Schut et al., 2017). The role of eye movements in the functioning of VWM goes beyond simply the resampling of visual information. Even when the external world does not provide the relevant visual information, studies still find that participants make eye movements to the locations of previously relevant visual information (e.g., Richardson & Spivey, 2000). This suggests that visual memoranda in VWM are linked to locations in the external world (Ferreira et al., 2008).

A complete understanding of visual memory requires an embodied approach (Gibson, 1979; Rowlands, 2010; Shapiro, 2011) which embraces the external visual world as an actual memory storage that should be incorporated in traditional memory models. We can recruit our environment for achieving our goals with a minimum expenditure of our scarce mental resources (Hutchins, 1995). Current memory models ignore an important property of the human brain: our brain is an energy-efficient system that aims to minimize its load. Instead of using the energy-consuming internal memory, our brain can rely on the external visual world to maintain important visual information.

The idea of offloading our cognitive capacity by engaging the outside world as a form of an external cognitive resource offers an interesting perspective from which we can investigate the nature of VWM. Philosophers that consider our mind as inseparable from the body and the environment we inhabit date back to Kant and Heidegger. The concept of an external memory is also in line with the extended mind thesis, which claims that a person's mind and associated cognitive processing extend into the external world and that cognition is the result of a continuous interaction between external and internal processes (Clark & Chalmers, 1998; Hutchins, 1995).

Traditional models of VWM strictly incorporate internal memory storage and ignore our interaction with the external visual world to maintain information.

This is surprising because it is already known that we frequently adopt the external world as a memory resource in our daily life. For instance, think of the use of the term "memory" for a mobile phone: we store information about phone numbers in the external memory that is our phone, safe and secure, with the only requirement that we remember *where* we have stored the information. Similarly, we write information down in an external memory to minimize the load on our internal memory system or remind ourselves not to forget our keys by placing them at a salient position in the external world. In my view, the same energy-efficient principle holds for the visual memory of the world around us. An embodied approach to visual working memory extends the definition of what is memory. Memory should refer to storage, irrespective of whether it is an internal or external memory storage.

How to measure?

If we indeed use the world as an external memory, there is a continuous decision about whether or not to store the visual features of an item internally or externally. This decision is then based on a trade-off between the costs associated with the execution of a saccade and the costs of storing the visual features of an object in visual working memory. Although the idea of a decision process between internal storage and external sampling is well grounded in the literature, less is known about the nature of the costs that determine this decision process. Below, I will discuss an example in which the trade-off is influenced by the "time costs" associated with the execution of an eye movement. Furthermore, costs might relate to the energy required to perform a specific movement (Gallivan et al., 2018), to perform a specific task (Alfandari et al., 2019; Droll & Hayhoe, 2007), or to the reward associated with a specific movement (Hayhoe & Ballard, 2005).

It automatically follows from the assumption of such a decision process that internal visual working memory will be used in those situations in which the costs of a saccade exceed a certain threshold. An example of a task to measure external memory use is the "copying task", in which the observer has to copy a complicated figure constituting simple coloured shapes ("the template") using the mouse. Previous research has shown that participants

perform many eye movements between the template and the workspace while performing a copying task, indicating that the template is not fully memorized in internal visual memory and that observers rely on the objects in the external world while only having to maintain the location of the objects in spatial working memory (Ballard et al., 1995). In this task, participants have the freedom to choose their own task parameters. For instance, they can choose not to rely on internal visual working memory and store the features of the template in external memory. Minimizing the use of internal visual working memory will therefore result in a high proportion of eye movements between the template and the workspace. In these types of tasks, observers frequently adopt a "just-in-time" strategy in which information is extracted from the external world only at the exact timing when that information is relevant for the current task (Triesch et al., 2003). In this situation, the reluctance to use internal visual working memory can be explained by the fact that such memory is more expensive to use compared to a strategy in which the information is stored in external memory. The conclusion that eye movements are an integral part of natural tasks (in which information is physically present) (Hayhoe & Ballard, 2005) is therefore crucial evidence for the view presented in the current paper.

In a recent study, we successfully observed a reluctance for the use of "expensive" memory when time costs associated with a saccade are increased, confirming the presence of a trade-off between storing information in VWM and making saccades (Somai et al., 2020). We influenced the trade-off between eye movements and VWM utilization by introducing a cost to a saccade. If there is an adaptive trade-off between using the external visual world and VWM, the trade-off should be influenced by increasing the cost associated with using external information. Higher costs were created by adding a delay in stimulus availability to a copying task (i.e., the time between the landing of the saccade on the template and the appearance of the template). By removing the template from the screen at the start of a saccade towards the template and delaying its presentation after the saccade, the (time) cost of a saccade is increased. This experiment included three different delays after which the template was revealed (250, 1500, and 3000 ms). Results showed that the increased saccade cost results in less saccades towards the

model and an increased dwell time on the model. These results suggest a shift from making eye movements towards taxing internal VWM. Our findings reveal that the trade-off between executing eye movements and building an internal representation of our world is based on an adaptive mechanism, governed by cost-efficiency (see also, Melnik et al., 2018).

These results were reminiscent of the findings by Droll and Hayhoe (2007). These authors manipulated the cognitive cost of the task by increasing the number of visual features or changing block sorting rules in an unpredictable fashion. The behavioural and eye-tracking results showed that the memory/ gaze trade-off is highly dynamic and heavily depends on the number of task-relevant features that need to be tracked as well as the unpredictability of the task. Specifically, if a task can be easily resolved with limited usage of working memory (i.e., at low capacity), then visual working memory is the preferred mechanism as it offsets the cost of fixating and re-fixating and externalizing mental computations, which itself carries a cognitive cost. Conversely, when visual working memory is taxed by a high number of features or the task is unpredictable, whereby encoding and maintenance of information is inefficient and inaccurate, participants rely on just-in-time saccades and re-fixations as their strategy of choice due to the lower cognitive cost. Droll and Hayhoe concluded that the assumption of object invariance also heavily dictates usage of visual working memory; if an object is simple and invariant and the visual task predictable, it will be encoded and executed in working memory as that is the least cognitively costly strategy.

Besides the costs associated with a saccade and the costs of storing the visual features of an object in visual working memory, the decision to sample information from the external world is also influenced by the uncertainty of one's own mental state. For instance, observers frequently made eye movements towards the cue for a subsequent search task, even when they had already learned the identity of the search cue over a course of trials (Alfandari et al., 2019). Rather than relying on their memory, observers chose to have another look at the search cue, perhaps in an effort to decrease the uncertainty associated with the accuracy of the internally stored search cue (i.e., "just to be sure"). From these studies, it is evident that eye movement patterns provide a unique insight into the current internal mental state of an observer.

There is already some research on the neural correlates of our interaction with objects still within view. As mentioned, previous research on VWM was conducted in situations in which stimuli disappear, mostly in the context of static, flat 2d shapes and objects disappearing from a computer screen. In a recent study, Tsubomi et al. (2013) examined neural responses when viewing targets that remained visible until response. They used the CDA component, an especially strong neural marker that corresponds to not only the amount of active representations maintained in the working memory, but also the rate of their decay and discarding (Vogel & Machizawa, 2004). When they compared the neural activation pattern with that found in a normal change detection task, they found that the two patterns were identical. This finding suggests that visual working "memory" is not a memory in the traditional sense of the word. Rather, working memory could be interpreted as a powerful but limited cognitive resource deployed according to visual task demands, regardless of the object's presence or absence.

Conclusion

Because the brain only stores a limited amount of visual information in VWM, there is a continuous decision about which information to internalize and which information to leave in the external world for (possible) access later in time. Current theories do not take into account the fact that our visual world is relatively stable, and that we may not need to store information when it remains externally available. Traditionally, experiments on VWM use visual stimuli that are presented briefly or change features rapidly and the participant's response is entirely contingent on their ability to use their working memory capacity effectively. This is the correct approach if we are mostly interested in the memory aspect of visual working memory; however, we have seen that the memory component of visual working memory is perhaps not its most crucial one. A more natural approach to visual tasks is required, with the ability to create an external and stable visual reference.

Although I have used the word "embodied" cognition here when discussing the world as an external memory (see also, Melnik et al., 2018), one could also argue that the term "embedded" cognition is perhaps more appropriate, as embedded cognition focuses on the strategies organisms use to off-load cognitive processing onto the environment (Hutchins, 1995), rather than on peripheral representations of memory. Given that embodied cognition is generally defined as the influence of the constraints of the body of the individual on cognition (e.g., Gallivan et al., 2018) and that I have proposed that the decision to internalize information in VWM is partly determined by the constraints of our oculomotor system, I prefer the term "embodied".

Our brain aims for an optimal balance between storing information in expensive and vulnerable VWM and leaving information in the outside world without internal storage (i.e., an internal mental economy). The outcome of this trade-off is determined by the costs of storage and the cost of (re-)acquiring the visual information in the external world. To fully understand human cognition, it is not enough to focus only on what goes on inside the skull, because our abilities are supported by our environment.

Disclosure statement

No potential conflict of interest was reported by the author(s).

References

Alfandari, D., Belopolsky, A. V., & Olivers, C. N. L. (2019). Eye movements reveal learning and information-seeking in attentional template acquisition. *Visual Cognition*, *27*(5–8), 467–486. https://doi.org/10.1080/13506285.2019.1636918

Baddeley, A. D., & Hitch, G. (1974). Working memory. In G. Bower (Ed.), *The psychology of learning and motivation* (Vol. VIII, pp. 47–89). Academic Press.

Ballard, D. H., Hayhoe, M. M., & Pelz, J. B. (1995). Memory representations in natural tasks. *Journal of Cognitive Neuroscience*, *7*(1), 66–80. https://doi.org/10.1162/jocn.1995.7.1.66

Berch, D. B., Krikorian, R., & Huha, E. M. (1998). The corsi block-tapping task: Methodological and theoretical considerations. *Brain and Cognition*, *38*(3), 317–338. https://doi.org/10.1006/brcg.1998.1039

Clark, A., & Chalmers, D. J. (1998). The extended mind. *Analysis*, *58*(1), 7–19. https://doi.org/10.1093/analys/58.1.7

Corbetta, M. (1998). Frontoparietal cortical networks for directing attention and the eye to visual locations: Identical, independent, or overlapping neural systems? *Proceedings of the National Academy of Sciences*, *95*(3), 831–838. https://doi.org/10.1073/pnas.95.3.831

Cowan, N. (2005). *Working memory capacity*. Psychology Press.

Deubel, H., & Schneider, W. X. (1996). Saccade target selection and object recognition: Evidence for a common attentional mechanism. *Vision Research*, *36*(12), 1827–1837. https://doi.org/10.1016/0042-6989(95)00294-4

Droll, J. A., & Hayhoe, M. M. (2007). Trade-offs between gaze and working memory use. *Journal of Experimental Psychology: Human Perception and Performance*, 33(6), 1352–1365. https://doi.org/10.1037/0096-1523.33.6.1352

Ferreira, F., Apel, J., & Henderson, J. M. (2008). Taking a new look at looking at nothing. *Trends in Cognitive Sciences*, 12(11), 405–410. https://doi.org/10.1016/j.tics.2008.07.007

Gallivan, J. P., Chapman, C. S., Wolpert, D. M., & Flanagan, J. R. (2018). Decision-making in sensorimotor control. *Nature Reviews Neuroscience*, 19(9), 519–534. https://doi.org/10.1038/s41583-018-0045-9

Gayet, S., Guggenmos, M., Christophel, T., Haynes, J.-D., Paffen, C. L., Van der Stigchel, S., & Sterzer, P. (2017). Visual working memory enhances the neural response to matching visual input. *The Journal of Neuroscience*, 37(28), 6638–6647. https://doi.org/10.1523/JNEUROSCI.3418-16.2017

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

Haber, R. N. (1983). The impending demise of the icon: A critique of the concept of iconic storage in visual information processing. *Behavioral and Brain Sciences*, 6(1), 1–11. https://doi.org/10.1017/S0140525X0001428X

Harrison, S. A., & Tong, F. (2009). Decoding reveals the contents of visual working memory in early visual areas. *Nature*, 458 (7238), 632–635. https://doi.org/10.1038/nature07832

Hayhoe, M., & Ballard, D. (2005). Eye movements in natural behavior. *Trends in Cognitive Sciences*, 9(4), 188–194. https://doi.org/10.1016/j.tics.2005.02.009

Hollingworth, A., & Luck, S. J. (2009). The role of visual working memory (VWM) in the control of gaze during visual search. *Attention, Perception & Psychophysics*, 71(4), 936–949. https://doi.org/10.3758/APP.71.4.936

Hutchins, E. (1995). *Cognition in the Wild*. MIT Press.

Kessels, R. P. C., van Zandvoort, M. J., Postma, A., Kappelle, L. J., & de Haan, E. H. F. (2000). The Corsi block-tapping task: Standardization and normative data. *Applied Neuropsychology*, 7(4), 252–258. https://doi.org/10.1207/S15324826AN0704_8

Logie, R. H. (1995). *Visuo-spatial working memory*. Lawrence Erlbaum.

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

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

MacKay, D. M. (1967). Ways of looking at perception. In W. Wathen-Dunn (Ed.), *Models for the perception of speech and visual form* (pp. 25–43). MIT Press.

Melnik, A., Schuler, F., Rothkopf, C. A., & Konig, P. (2018). The world as an external memory: The price of saccades in a sensorimotor task. *Frontiers in Behavioral Neuroscience*, 12 (253). https://doi.org/10.3389/fnbeh.2018.00253

O'Regan, J. K. (1992). Solving the "real" mysteries of visual perception: The world as an outside memory.. *Canadian Journal of Psychology*, 46(3), 461–488. https://doi.org/10.1037/h0084327

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

Rensink, R. A., O'Regan, J. K., & Clark, J. J. (1997). To see or not to see: The need for attention to perceive changes in scenes. *Psychological Science*, 8(5), 368–373. https://doi.org/10.1111/j.1467-9280.1997.tb00427.x

Richardson, D. C., & Spivey, M. J. (2000). Representation, space and Hollywood Squares: Looking at things that aren't there anymore. *Cognition*, 76(3), 269–295. https://doi.org/10.1016/S0010-0277(00)00084-6

Rowlands, M. (2010). The mind embedded. In *The new science of the mind: From extended mind to embodied phenomenology*. MIT Press.

Schut, M. J., Van der Stoep, N., & Van der Stigchel, S. (2017). The cost of making an eye movement: A direct link between visual working memory and saccade execution. *Journal of Vision*, 17(15). https://doi.org/10.1167/17.6.15

Shapiro, L. (2011). *Embodied cognition*. Routledge.

Simons, D. J., & Levin, D. T. (1997). Change blindness. *Trends in Cognitive Sciences*, 1(7), 261–267. https://doi.org/10.1016/S1364-6613(97)01080-2

Somai, R., Schut, M. J., & Van der Stigchel, S. (2020). Evidence for the world as an external memory: A trade-off between internal and external visual memory storage. *Cortex*, 122, 108–114. https://doi.org/10.1016/j.cortex.2018.12.017

Theeuwes, J. (2012). Automatic control of visual selection. In M. Dodd & J. Flowers (Eds.), *The influence of attention, learning, and motivation on visual search* (pp. 23–62). Springer.

Triesch, J., Ballard, D. H., Hayhoe, M. M., & Sullivan, B. T. (2003). What you see is what you need. *Journal of Vision*, 3(1), 86–94. https://doi.org/10.1167/3.1.9

Tsubomi, H., Fukuda, K., Watanabe, K., & Vogel, E. K. (2013). Neural limits to representing objects still within view. *Journal of Neuroscience*, 33(19), 8257–8263. https://doi.org/10.1523/JNEUROSCI.5348-12.2013

Turvey, M. T. (1977). Contrasting orientations to the theory of visual information processing. *Psychological Review*, 84(1), 67–88. https://doi.org/10.1037/0033-295X.84.1.67

Van der Stigchel, S., & Hollingworth, A. (2018). Visuospatial working memory as a fundamental component of the eye movement system. *Current Directions in Psychological Science*, 27(2), 136–143. https://doi.org/10.1177/0963721417741710

Van der Stigchel, S., & Theeuwes, J. (2005). The influence of attending to multiple locations on eye movements. *Vision Research*, 45 (15), 1921–1927. https://doi.org/10.1016/j.visres.2005.02.002

Van der Stigchel, S., & Theeuwes, J. (2007). The relationship between covert and overt attention in endogenous cueing. *Perception & Psychophysics*, 69(5), 719–731. https://doi.org/10.3758/BF03193774

Van Zoest, W., Van der Stigchel, S., & Donk, M. (2017). Conditional control in visual selection. *Attention, Perception & Psychophysics*, 79(6), 1555–1572. https://doi.org/10.3758/s13414-017-1352-3

Vogel, E. K., & Machizawa, M. G. (2004). Neural activity predicts individual differences in visual working memory capacity. *Nature*, 428(6984), 748–751. https://doi.org/10.1038/nature02447