
Downloaded from:

Usage Guidelines:
Please refer to usage guidelines at contact lib-eprints@bbk.ac.uk. or alternatively

© 2019 The British Psychological Society. This article may not exactly replicate the final version published. It is not the copy of record.

Current directions in visual working memory research: an introduction and emerging insights

Heinrich René Liesefeld¹,², & Hermann J. Müller¹,³

¹Department Psychologie, Ludwig-Maximilians-Universität, München, Germany; ²Graduate School of Systemic Neurosciences, Ludwig-Maximilians-Universität München, Germany; ³Department of Psychological Sciences, Birkbeck College, University of London, UK
Abstract

Visual working memory (VWM) is a core construct in the cognitive (neuro-)sciences, assumed to serve as a hub for information exchange and thus supporting a multitude of cognitive functions related to processing visual information. Here, we give an introduction into key terms and paradigms and an overview of ongoing debates in the field, to which the articles collected in this Special Issue on “Current Directions in Visual Working Memory Research” contribute. Our aim is to extract, from this overview, some ‘emerging’ theoretical insights concerning questions such as the optimal way to examine VWM, which types of mental representations contribute to performance on VMW tasks, and how VWM keeps features from the same object together and apart from features of concurrently maintained objects (the binding problem).

Keywords: visual cognition, working memory, attention, binding, interference
Current directions in visual working memory research: an introduction and emerging insights

Visual working memory (VWM) – the mental workspace for the temporary storage and manipulation of visual information – is a central cognitive faculty at the interface between visual perception and goal-directed action. It is pivotal for making sense of the constantly changing stream of information arriving through the visual system: Objects move, our eyes move, our body moves, even the whole world may appear to move (such as on a merry-go-round). To keep track of agents and things in this rapid flow, it is crucial to maintain and match representations of relevant visual objects across all these changes (Aagten-Murphy & Bays, 2019; Cronin & Irwin, 2018). But beyond holding visual information passively (a function implied by the term ‘memory’), VWM is thought to be a general-purpose visual representation system that is involved in a multitude of cognitive functions requiring the (active) processing of currently available visual information in order to achieve the goals of the task at hand (Luck & Vogel, 2013; Zimmer, 2008). Reflecting its importance as a central cognitive hub and its relation to many crucial cognitive functions (Fukuda, Vogel, Mayr, & Awh, 2010; Unsworth, Fukuda, Awh, & Vogel, 2014), VWM has developed, over the past decades, into a core topic in the cognitive (neuro-)sciences, sparking an abundance of theoretical and methodological advances that range from general models of cognition and consciousness (Block, 2011; Bundesen, Habekost, & Kyllingsbaek, 2005, 2011; Dehaene, & Naccache, 2001; Schneider, 2013; Soto & Silvanto, 2016; Usher, Bronfman, Talmor, Jacobson, & Eitam, 2018) to the trial-wise decoding of specific mental contents from neuroimaging data (brain reading; Christophel, Klink, Spitzer, Roelfsema, & Haynes, 2017; Harrison & Tong, 2009; Serences, Ester, Vogel, & Awh, 2009; Sprague, Ester, & Serences, 2014).

On this background, in July 2017, we hosted an international symposium on VWM at Ludwig-Maximilians-University Munich, which aimed to bring together a select group of experts representing different theoretical stances and methodological approaches in this rapidly growing field. Our intention was to foster scientific exchange and discussions across the various viewpoints and help identify the most urgent and promising questions for future research. The result did meet, if not exceed, our high
expectations. To make (at least some of) the spirit of this symposium available to a broader audience, we decided to edit the present special issue, featuring work from many participants of the symposium as well as a select number of experts who could not attend at the time. What we, hopefully, have achieved is a representative, state-of-the-art overview of the field, spanning many of the various viewpoints and currently debated issues.

To put the collection of articles into context, in particular (though not only) for readers less familiar with the vibrating research on VWM, we go on to provide a selective overview on the history of the field and the experimental paradigms and theoretical issues that drove and continue to drive much of the research. As detailed below, all articles in the present issue, in one way or another, relate to these various lines of ideas and the respective debates.

Most strands of psychological investigation start off with beautiful simplicity, but become more complex and seemingly ‘unwieldy’ as research progresses. VWM research is no exception to this. Originally, there was broad consensus that VWM could hold some four bound objects stored in separate slots (Luck & Vogel, 1997), that the usage of VWM slots could be accurately tracked via neuronal markers (Todd & Marois, 2004; Vogel & Machizawa, 2004), and that performance on the standardly used change-detection task (Pashler, 1988) purely reflects VWM functioning, free of any extraneous factors (Cowan, 2001).

In the (one-shot, whole-display) change-detection task, several objects (most often: coloured squares) are briefly presented in a first (memory) display and after a short retention interval (typically ~1 s), a second (test) display is shown. The second display is identical to the first in one half of randomly intermixed trials, but in the other half one object has changed (typically its colour). Participants’ task is to decide whether a change was present or not. A simple formula can be applied to calculate, from the observed error pattern, the number of objects a given participant can maintain in VWM \( k \) (Pashler, 1988; Rouder, Morey, Morey, & Cowan, 2011). Performance on this task was – in contrast to previous, largely verbal, working-memory research – considered to be free from confounding influences such as long-term memory and response interference, and, as all information was available in one ‘snapshot’, there was no
reason to treat the various objects differently (Cowan, 2001). However, subsequent research inspired by these initial findings resulted in the accumulation of ever more data that was not easily reconcilable with these initial assumptions and the original, simple theory.

For instance, the assumption that bound objects are held in VWM was supported by a series of carefully executed change-detection experiments showing that any number of features could be stored for each object with no additional costs for additional features (Luck & Vogel, 1997). This strong assumption was soon challenged by studies that observed a clear decrease in performance as the number of features per object increased. First, costs were shown for features from the same feature dimension (e.g., Olson & Jiang, 2002; Wheeler & Treisman, 2002), for increasing object complexity (Alvarez & Cavanagh, 2004), and for assigning features to different parts of the same object (Xu, 2002). Further research revealed costs even for features from different dimensions on the same simple-object part (Fougnie, Asplund, & Marois, 2010). The broadest consensus at present appears to be that, although there is some object benefit (i.e., more information can be stored when it is distributed across fewer objects, likely due to the features of an object sharing the same location), storing additional features is not cost-free (Fougnie, Cormiea, & Alvarez, 2013).

Whether costly or not, the brain often has to code multiple features of an object and, additionally, information on which features belong to which object (Treisman, 1996; Treisman & Gelade, 1980). This binding problem is a recurrent theme in working-memory research and the present special issue. In their contribution, Schneegans and Bays (2019) give a comprehensive overview of the binding problem in visual working memory and how this problem might be solved on the neuronal level. They conclude that both object benefits and costs of storing additional features can be explained by noisy, distributed VWM representations of features, and that it is unnecessary to postulate a separate resource for binding features within an object. Instead, features might be bound by default via their common position in space (Schneegans & Bays, 2017) or time (most likely, ordinal position; Oberauer, Lewandowsky, Farrell, Jarrold, & Greaves, 2012; Swan & Wyble, 2014). Whether space or time is used for binding features is not easily distinguishable: when objects are presented concurrently at different positions, space can be
remapped onto time by sequential attention allocations (Zimmer & Liesefeld, 2011); and when objects are presented sequentially at the same location, time can be remapped onto space by internally assigning objects to different unoccupied locations (Abrahamse, Van Dijck, Majerus, & Fias, 2014; see also Oberauer, 2009). Another hotly debated question concerns whether VWM capacity is actually limited by slots or rather by a continuous, infinitely divisible resource (Bays & Husain, 2008). Schneegans and Bays favour the latter assumption and explain this limit in terms of representational noise (see also Oberauer & Lin, 2017; Wilken & Ma, 2004)

The slots-vs.-resource debate was largely based on data from a new procedure that can provide a more fine-grained picture of VWM performance than the classical change-detection task: Instead of comparing the memory display to a second test display, the continuous-report task requires participants to reproduce an object feature chosen from a (virtually) continuous circular space (e.g., by choosing the colour of a probed object from a colour wheel; Wilken & Ma, 2004; Zhang & Luck, 2008; or by adjusting a probe object, e.g., Bays, Wu, & Husain, 2011). This task is also referred to as analogue/cued-recall or delayed-estimation task. In contrast to the change-detection task, the continuous-report task does not provide a binary outcome on each trial (correct vs. incorrect response), but a (virtually) continuous measure: the distance between the correct feature value and the actual response (estimation error). Zhang and Luck (2008) proposed to analyse data in this task by fitting a mixture of a van Mises (the circular equivalent of a Gaussian) and a uniform distribution to the distribution of estimation errors and interpreting this model’s parameter values. This mixture model neatly fits the typically observed error distribution and has since been ubiquitously applied to continuous-report data. However, several arguments against the use of the mixture model have been raised (see, in particular, Ma, 2019), so that it is probably advisable to analyse raw estimation errors (in addition) or to apply more elaborate, mechanistic models of performance on the continuous-report task (e.g., Oberauer & Lin, 2017; Schneegans & Bays, 2017; Swan & Wyble, 2004).

Various neuronal correlates of VWM representations have been proposed. Early research identified activity in/over parietal areas as measured by EEG contralateral to the remembered items (contralateral
delay activity, CDA; Vogel & Machizawa, 2004; for a recent review, see Luria, Balaban, Awh, & Vogel, 2016) or by fMRI (Todd & Marois, 2004; Xu & Chun, 2006) as a potential neuronal correlate of VWM maintenance. Subsequent studies decoded VWM content from early visual areas (Harrison & Tong, 2009; Serences et al., 2009). It turned out that decoding VWM content is possible from many areas throughout the human brain (for a recent review, see Christophel et al., 2017). As concurrent and partially overlapping activation of several neuronal populations would make it impossible to decide which features belong to which objects (the binding problem), it was speculated that VWM-related parietal activity reflects a pointer system, with each pointer connecting all features of one object. Indeed, the parietal activity does not appear to be specific to the maintenance of visual information after it is gone; rather, it also reflects processing of physically available information in VWM (Drew & Vogel, 2008; Tsubomi, Fukuda, Watanabe, & Vogel, 2013).

As another – neuronally inspired – potential solution to the binding problem, Lisman and colleagues proposed that objects are differentiated in time by differential temporal spiking patterns (Lisman & Jensen, 2013; Lisman & Idiart, 1995; but see Schneegans & Bays, 2019). Neurons coding a certain feature fire at a high frequency (gamma, i.e. > 30 Hz), with different representations (gamma cycles) being locked to different phases of a theta cycle (i.e., 4-7 Hz). In other words, each item in VWM is repeatedly reactivated (gamma) in a given order (theta). Accordingly, the coupling of gamma and theta phase has been explored as a neuronal marker of VWM maintenance (Sauseng et al., 2009), and this idea was used to explain capacity limits and the (purported) slot-like fashion of VWM storage: VWM capacity is limited by the number of complete gamma waves that can be nested within a theta cycle of a given length.

Alternatively, according to a more recent proposal, items are coded as gamma bursts (i.e., multiple cycles; Herman, Lundqvist, & Lansner, 2013; Van Vugt, Chakravarthi, & Lachaux, 2014) and only one item is coded per theta period, with different items being refreshed in consecutive theta cycles. During a theta cycle, gamma amplitude first increases and then decreases, yielding a coupling of gamma amplitude to theta phase. In their contribution to the special issue, Sauseng, Peylo, Biel, Friedrich and Romberg-Taylor (2019) review the pertinent cross-frequency-coupling literature and suggest that the more recent version of
this theory (Herman et al., 2013) and a newly proposed alternative are compatible with continuous-
resource as well as fixed-slots conceptions of VWM capacity. In particular, they argue that some
flexibility in the length of theta and/or gamma cycles yields flexibility in the number and precision of
VWM representations.

Much research has shown that not all VWM representations are created equal, and that
representational quality varies considerably across trials and objects (Fougnie, Suchow, & Alvarez, 2012;
vanden Berg, Shin, Chou, George, & Ma, 2012). In their contribution to the special issue, Czoschke,
Fischer, Beitner, Kaiser, and Bledowski (2019) examine two sources of this variability: processing objects
on the previous trial(s) and processing other objects on the current trial. With a sequential presentation of
two objects per trial, they observe that VWM representations are attracted towards the last item on the
previous trial and repulsed away from the preceding item on the same trial. These effects might reflect
mechanisms contributing to the separation and integration of information across objects and episodes.

In their contribution to the special issue, Bae and Luck (2019) examine how representing a single
item in VWM is affected by performing a simple task during maintenance. They combine the continuous-
report task with the classical dual-task interference paradigm that dominated the working-memory
literature before research on VWM gained momentum (Baddeley, 1986; Logie, 1995). In this paradigm,
participants have to perform a secondary task during working-memory maintenance, which is supposed to
either draw or not draw on the same working-memory system or component. Bae and Luck’s results show
that processing a letter deteriorates the representation of a single orientation, in that it becomes less precise
and biased more strongly away from cardinal orientations (see also Hardman, Vergauwe, & Ricker, 2017).
No such effect is found for processing an auditory stimulus during single-orientation maintenance. This
dissociation indicates that processing information from the same modality displaces the VWM object from
a special state (‘focus of attention’, see Oberauer, 2002; Oberauer & Lin, 2017), leaving only a relatively
coarse representation. Having to remember only one item on each trial is critical in this study, because this
rules out interference from other items on the same trial. Bae and Luck’s study also nicely demonstrates a
major strength of the continuous-report task: it provides a particularly sensitive measure of subtle effects
on VWM performance. Remembering only one object would produce ceiling performance in the change-detection task, making it impossible to examine effects on the representation of a single item.

In their contribution to the special issue, Tabi, Husain and Manohar (2019) demonstrate a drawback of a standard version of the continuous-report task, which requires participants to adjust a memory probe: performance improves if the probe object is initially set to neutral with respect to the probed feature, instead of taking a random feature value as is common practice in this paradigm. Thus, processing the random probe feature likely partially overwrites the probed memory representation. A similar interference is likely induced by the commonly employed colour wheel (Souza, Rerko, & Oberauer, 2016). Tabi et al.’s new method to start off with a neutral probe is designed to eliminate this source of interference by allowing participants to first transfer the probed feature into a more robust state (see Souza et al., 2016, for a related effect of retro-cues, i.e., cues given during the retention period before the test).

Probably, such interference does not occur in the change-detection task. One potential interpretation of the whole-display change-detection task is that the onset of the test display produces a spatial representation akin to a saliency map in visual search that (noisily) encodes the mismatch with the memorized display, with registration of a mismatch signal being used to decide upon the presence of a change (Hyun, Woodman, Vogel, Hollingworth, & Luck, 2009; Liesefeld, Liesefeld, Müller, & Rangelov, 2017). On that account, the interference by the test display would occur only after the critical mismatch signal is already produced, so that the change-detection task would indeed provide a relatively pure (albeit coarse) picture of VWM performance (but see, Alvarez & Thompson, 2009).

A further advantage of the change-detection task is that, in addition to accuracy, it allows reaction time to be examined as a measure of the speed of comparing VWM entries with visual input (Donkin, Nosofsky, Gold, & Shiffrin, 2013; Gilchrist & Cowan; Hyun et al., 2009; Liesefeld et al., 2017; Park, Zhang, & Hyun, 2017). The interpretation of reaction times is less straight-forward in the continuous-report task, because it takes quite some time until a response is selected and executed (moving the arrow to the position on the colour wheel or toggling through all values with the probe), allowing for all kinds of confounding influences and measurement noise (see, e.g., Tabi et al., 2019). Thus, the two tasks have
complementary strengths and weaknesses and can provide different perspectives on VWM performance, and comprehensive models of VWM should explain performance on both (and other, e.g., localization, Liesefeld, Liesefeld, & Müller, 2019; van den Berg et al., 2012) tasks.

An (unwanted) source of variability in all types of VWM tasks arises from the possibility that people do (at times) not only represent individual objects, but (additionally) what several objects in the display have in common. This might include maintaining chunks of information (Miller, 1957; Morey, Cong, Zheng, Price, & Morey, 2015; Peterson & Berryhill, 2013; Quinlan & Cohen, 2012) or ensemble representations (Brady & Alvarez, 2011, 2015a,b; Brady & Tenenbaum, 2013). In the classical colour change-detection task, repetition of colours within a display provides opportunities for such compressed representations. In her contribution to the special issue, Morey (2019) directly manipulated the potential for compression by presenting up to three objects of the same colour. She observed increased detection rates for these redundant objects and also – in young, but less so in older adults – a spill-over of VWM capacity to objects with unique colours. Furthermore, using pupil dilation as a marker of mental effort, she found decreased effort with increasing redundancy in young participants.

In our contribution to the special issue (Liesefeld et al., 2019), we tried to minimize opportunities for compressed coding by explicitly avoiding colour repetitions within a display. Still, even in this purest version of the standard colour change-detection task, we find indications for the involvement of ensemble representations. In particular, we additionally asked participants to localize the change and compared localization performance to change-detection performance using mathematical models. The model providing the best account of our data indicates that VWM capacity is routinely overestimated when considering only change-detection performance and that VWM can actually hold only about 2.3 individual items, which contrasts with earlier estimates of 7±2 (Miller, 1957) or, respectively, 4±1 (Cowan, 2001) items. Given that a privileged, focal-attentional state (Bae & Luck, 2019; Oberauer, 2002; Oberauer & Lin, 2017) holds 1 of these 2.3 items, the capacity estimate for ‘usual’ VWM would effectively shrink to (a little over) 1 item.
Speculations about multiple, qualitatively different VWM states (see also Cowan, 1999; Oberauer, 2002) prompt another potential solution to the binding problem (or rather its dissolution). If ‘usual’ VWM indeed can hold only one item, one item is in the focus of attention, and any additional information is coded as an ensemble, one needs to assume only one representation in each state to explain typical estimates of VWM capacity of around 3 items. Consequently, there is no further need to keep features from an object bound together and apart from other objects. The notion that a multitude of qualitatively different representations, at different levels of abstraction, contribute to performance on VWM tasks would be in line with recent neuroimaging evidence for VWM representations in several, distributed brain areas (Christophel et al., 2017). The ever-increasing list of contributors to performance on VWM tasks also hints at a general “problem” in cognitive science: humans are highly adaptive and exploit whatever cognitive resources they have at their disposal to solve a given (laboratory) task without regard to the (by the scientist) targeted cognitive construct and the aesthetically pleasing simplicity of cognitive theories (see also Zimmer & Liesefeld, 2011).

Not only does attention affect VWM performance (by raising one item to a special state; Griffin & Nobre, 2003; Oberauer, 2002; Oberauer & Lin, 2017; Souza & Oberauer, 2016), but VWM is also thought to influence attentional dynamics. One common assumption is that an item in the special state operates as a search template (Duncan & Humphreys, 1989; Hollingworth & Hwang, 2013; Olivers, Peters, Houtkamp, & Roelfsema, 2011; Schneider, 2013). In their contribution to the special issue, Berggren and Eimer (2019) examine how storing shapes in VWM conflicts with spatial templates (i.e., maintenance of potential target locations). Using an electrophysiological marker of attention allocations (N2pc; Eimer, 1996; Luck & Hillyard, 1994a,b), they observe a delay in attention allocation towards the target when VWM load is high, indicating that VWM for shapes and spatial templates share a common resource.

Working memory also appears to be involved in the familiarization with visual-search scenes. It has been shown that repeating spatial stimulus configurations (‘context’) speeds up the search due to incidental long-term learning of spatial distractor-to-target relations (contextual cueing). In his contribution, Pollmann (2019) reviews the literature on the involvement of spatial working memory in the
contextual-cueing effect. Various studies have shown that spatial working-memory load can disrupt the effect. This is, however, true only if applied during the test phase of the task, indicating that spatial working-memory load does not impede the acquisition of contextual information; rather, the retrieval of this information from long-term memory operates via working memory and may thus be hampered by spatial working-memory load. In their contribution, Annac, Zang, Müller and Geyer (2019) report that this is not necessarily the case: the working-memory bottleneck can be bypassed as a result of training. That is, after an extended period of training, the contextual-cueing effect was found to emerge even with concurrent working-memory load during test. Surprisingly, Annac et al. even observed an amplification of the contextual-cueing effect by concurrent working-memory load, indicating that, under these conditions, context information may be directly retrieved from long-term memory (see Reinhart, McLenahan, & Woodman, 2016, for a related finding).

Another core topic in VWM research are interindividual differences in VWM capacity and whether capacity is fixed or can be boosted (Fukuda et al., 2010; Liesefeld, Liesefeld, & Zimmer, 2014; Unsworth et al., 2014; Vogel & Machizawa, 2005; Vogel et al., 2005). Sauseng et al.’s (2019) review discussed above summarizes evidence that entrainment of certain theta frequencies via electrical cortical stimulation can increase (slower theta) or decrease (higher theta) VWM capacity (Vosskuhl, Huster, & Hermann, 2015; Wolinski, Cooper, Sauseng, & Romei, 2018). But we don’t need to turn to cortical stimulation to improve VWM performance. In their contribution to the special issue, Dodwell, Müller and Töllner (2019) examine how moderate aerobic exercise (vs. passive control) and body posture (upright on a treadmill vs. seated on a bike) while performing a speeded VWM task influence separable VWM processes. Exercise and upright posture speeded reaction times compared to passive (resting) and seated conditions, respectively. Examination of electrophysiological markers revealed that this reaction-time advantage was not due to faster VWM access, but rather to expedited processes of selecting the appropriate response to the target representation held in VWM.

Finally, the relation between stimulus material and the observer, such as familiarity (Curby & Gauthier, 2007; Curby, Glazek, & Gauthier, 2009; Jackson & Raymond, 2008) and emotional content
(Jackson, Linden, & Raymond, 2014; Jackson, Wu, Linden, & Raymond, 2009; Kensinger & Corkin, 2003; Maran, Sachse, & Furtner, 2015; Mather, 2007), also influences VWM performance. In their contribution, Curby et al. (2019) report a detrimental effect of fearful faces on VWM capacity that is consistent across four experiments. This is particularly notable, as earlier research had more often found VWM benefits, rather than costs, for emotional (including fearful, angry, sad, and happy) faces (e.g., Jackson et al., 2009, 2014; Sessa, Luria, Gotler, Jolicoeur, & Dell'acqua, 2011). Curby et al. test and discuss in great detail how effects of emotion on VWM performance are modulated by subtle differences in task designs (such as encoding duration and set size) and/or by the processing and maintenance of different emotions.

The collection of articles is rounded off by an integrative commentary by Rhodes and Cowan (2019) that we have not yet received at the time of writing this introduction. It will be of interest to the reader to compare the overlap of our introduction with their commentary.

Although some current topics, such as the active manipulation of VWM content (e.g., Christophel, Cichy, Hebart, & Haynes, 2015; Liesefeld & Zimmer, 2013) and the relationship of LTM and VWM (e.g., Brady, Störmer, & Alvarez, 2016; Schurgin, Cunningham, Egeth, & Brady, 2018) are not represented here, we believe that reading through the articles in this issue provides a relatively comprehensive overview on the current state of VWM research. In any case, we are proud to present this collection of articles to the readers of the *British Journal of Psychology* and thank the editor-in-chief, Stefan Schweinberger, for giving us the opportunity to compile this Special Issue. We hope that you all enjoy reading it as much as we enjoyed putting it together!

References


Bae, G.-Y., & Luck, S. J. (2019). What happens to an individual visual working memory representation when it is interrupted? *British Journal of Psychology*. Advance online publication. doi:10.1111/bjop.12339


INTRODUCTION TO VISUAL WORKING MEMORY


