Working memory and consciousness: The current state of play

Frontiers in Human Neuroscience 12:78

Marjan Persuh[1, #], Eric LaRock[2], and Jacob Berger[3]

[1] Department of Social Sciences, Human Services and Criminal Justice
Borough of Manhattan Community College
City University of New York
199 Chambers Street
New York, NY 10007

[2] Department of Philosophy
Oakland University
751 Mathematics and Science Center
Rochester, MI 48309

[3] Department of English and Philosophy
Idaho State University
921 South 8th Avenue, Stop 8056
Pocatello, ID 83209-8056

[#]Corresponding author

Marjan Persuh
Department of Social Sciences, Human Services and Criminal Justice
Borough of Manhattan Community College
City University of New York
199 Chambers Street
New York, NY 10007

Phone: +1 (212) 220-5226
Email: mpersuh@bmcc.cuny.edu

Abstract

Working memory, an important posit in cognitive science, allows one to temporarily store and manipulate information in the service of ongoing tasks. Working memory has been traditionally classified as an explicit memory system – that is, as operating on and maintaining only consciously perceived information. Recently, however, several studies have questioned this assumption, purporting to provide evidence for unconscious working memory. In this paper, we focus on visual working memory and critically examine these studies as well as studies of unconscious perception that seem to provide indirect evidence for unconscious working memory. Our analysis indicates that current evidence does not support an unconscious working memory store, though we offer independent reasons to think that working memory may operate on unconsciously perceived information.

Keywords: visual working memory; consciousness; unconscious perception; visual perception; visual awareness

1. Introduction

Working memory (WM), the capacity to temporarily store and manipulate information in the service of ongoing tasks (Baddeley, 1986), has been correlated with an array of cognitive abilities, including text comprehension, analytical thinking, and general intelligence (Fukuda, Vogel, Mayr, & Awh, 2010; Johnson et al., 2013). WM, and especially visual working memory (VWM), has also traditionally been linked to perceptual consciousness – that is, it is often assumed to operate on and maintain only consciously perceived information (Baddeley, 1986; Carruthers, 2015; Prinz, 2012).

The goal of this paper is to explore whether or not the contents of VWM are invariably conscious. Though some recent studies purport to demonstrate unconscious VWM (Bergstrom & Eriksson, 2014; Soto, Mantyla, & Silvanto, 2011), these results have been variously challenged (Carruthers, 2015; Prinz, 2012; Stein, Kaiser, & Hesselmann, 2016). We explore here many of these critiques, as well as several studies not previously discussed, often pursuing different lines of response. Though our analysis likewise indicates that current evidence does not support unconscious VWM, we offer independent reasons to think that WM may operate on unconsciously perceived information.

2. Working Memory

A major complicating factor in the debate over the existence of unconscious VWM is that there remains much uncertainty about how best to model the phenomenon of WM, consciously or otherwise. Thus we begin in this section by briefly exploring current models of WM and VWM.

2.1. Current Models of WM

Like both short-term memory (STM) and long-term memory (LTM) (Atkinson & Shiffrin, 1968), WM is typically characterized in terms of its functionality. Following the majority of the psychological literature, we define 'WM' as the storage system responsible for the maintenance of information in the service of ongoing work – that is, the system that makes available stored information for task-based manipulation – without imposing a limit on its duration or relationship to LTM (Baddeley, 1986; Luck & Vogel, 2013; Miller, Galanter, & Pribram, 1960).

Though perhaps the most influential account of WM is the multicomponent model proposed by Baddeley and Hitch (1974) – which includes two storage systems (a phonological loop and visuospatial sketchpad), a central executive, and more recently an episodic buffer (Baddeley, 2000) – this view has been slowly superseded by more recent state-based models (reviewed in Larocque, Lewis-Peacock, & Postle, 2014). Rather than postulating the existence of different systems (buffers) for different memory components, state-based models propose that attention to internal representations such as sensory, motoric, or LTM representations results in different states of information activation. State-based cognitive models have received much experimental support from contemporary cognitive neuroscience.

Cowan (1995), for example, proposes that information in WM exists in one of two states: a capacity-limited state, the so-called 'focus of attention' (FoA), or in capacity-unlimited state called 'activated LTM', which shows temporal decay (see also McElree, 1998; Oberauer, 2002). Such models have been developed to address a set of behavioral findings. For example, Oberauer and colleagues (Oberauer, 2001, 2002) used a retro-cue during a delay period to indicate relevant items from the memory set for the upcoming task. Cued items received attentional prioritization (FoA), whereas uncued items, which were not forgotten, were presumably stored in activated LTM. State-based models dubbed 'sensory' or 'sensorimotor-recruitment' models have also been developed for perceptual stimuli (Awh & Jonides, 2001; Magnussen, 2000).

Since arguably the most emphasized characteristic of WM is its storage-capacity limit, much work has focused on this aspect of the phenomenon. Two of the most widely cited studies are Miller's (1956) and Cowan's (2001), who reported an average capacity of seven items and four items, respectively, for verbal WM. In the visual domain, Luck and Vogel (1997) reported average capacity of around four individual objects. According to so-called 'slot models' of VWM, individual items are stored in a limited number of slots, whereas other items are discarded (Luck & Vogel, 1997). Continuous-resource models, by contrast, treat VWM as highly limited in capacity while allowing the distribution of resources among all items (Ma, Husain, & Bays, 2014). According to these models, the number of items remembered is not a fundamental metric, but rather the precision (quality) of memory. A recent variable-precision model further suggests that VWM precision varies from trial to trial and from item to item (Fougnie, Suchow, & Alvarez, 2012; van den Berg, Shin, Chou, George, & Ma, 2012).

Although it is clear that VWM is limited in capacity, there is currently no agreement about the nature of these limits. Several authors have demonstrated how slot and resource approaches could blend into one another (Souza, Rerko, Lin, & Oberauer, 2014); it seems likely that a final model, firmly grounded in neural data, will involve aspects of both slot and continuous-resource models (Wolfe, 2014).

In addition, much recent experimentation has attempted to identify the neural underpinnings on WM. Since the discovery of the persistent neuronal activity in monkey prefrontal cortex (PFC) during the delay interval of a WM task (Fuster & Alexander, 1971; Kubota & Niki, 1971) and related findings in human PFC with fMRI (Courtney, Ungerleider, Keil, & Haxby, 1997; Zarahn, Aguirre, & D'Esposito, 1997), it was widely believed that such activity reflects maintenance of WM representations.

This interpretation was, however, questioned when two prominent studies showed that stimulus information during delay periods can be decoded from primary visual cortex with multivariate pattern analysis (MVPA) of fMRI data in the absence of elevated signal levels (Harrison & Tong, 2009; Serences, Ester, Vogel, & Awh, 2009). Furthermore, by using a multiple step retro-cue design to specify the relevant items in a WM task, Lewis-Peacock and colleagues (Lewis-Peacock, Drysdale, Oberauer, & Postle, 2012) showed that MVPA evidence for the non-cued item dropped to the baseline, even though the item could be retrieved by a second retro-cue. These results suggest that persistent neural activity is not necessary to maintain item representations in WM. One prominent idea is that representations are sustained by modification of synaptic weights (Mongillo, Barak, & Tsodyks, 2008). A recent study provided converging evidence by using TMS instead of a second retro-cue, to activate memory for the non-cued item (Rose et al., 2016).

Experimental evidence suggests that several neural mechanisms, from intracellular to network based, contribute to WM (for an excellent review see D'Esposito & Postle, 2015). These findings support state-based models, and eliminate the need for specialized buffers. It has been suggested that because persistent neural activity or modulation of synaptic weights is likely a property of most neurons; WM representations arguably can be encoded by neuronal networks virtually anywhere in the brain (D'Esposito & Postle, 2015).

2.2. Varieties of Visual Short-Term Memory

A related obstacle to the study of unconscious VWM is the difficulty in distinguishing its operation from the operation of other visual short-term memory (VSTM) stores. In a well-known study that employed partial report, Sperling (1960) demonstrated the existence of a high-capacity, but limited-duration memory store that he termed 'iconic memory'. When post cued, participants were able to report letters from any row of a multi-row letter display. Although this memory store has a high capacity, it decays rapidly on the order of hundreds of milliseconds. According to the classical view, only a few items, selected from iconic memory by attentional mechanisms, form more durable and robust representations that last for several seconds, constituting VWM.

More recently, another type of VSTM has been proposed, so-called 'fragile VSTM' (Sligte, Scholte, & Lamme, 2008; Sligte, Vandenbroucke, Scholte, & Lamme, 2010), which purportedly has a lower capacity than iconic memory, retains high-resolution representations, and decays linearly over several (~ four) seconds. According to this proposal, VSTM consists of two limited-duration systems, iconic memory and fragile VSTM, which store many high-resolution representations. These are distinguished from the more robust and durable VWM, which has no duration parameters and stores only one or few high-resolution representations. The existence of fragile VSTM as opposed to mere iconic memory, however, remains controversial (Matsukura & Hollingworth, 2011).

Despite the debate over the fundamental nature of VWM and how it differs from other memory stores, we nonetheless believe that it is possible to assess the current state of evidence for and the possibility of unconscious VWM. In the next section, we explore some of the reasons that theorists have assumed that the contents of VWM are invariably conscious and offer reasons to think that this assumption is questionable.

3. Unconscious VWM

Though it is typically assumed that the contents of VWM are always or even must be conscious, the idea that VSTM systems can store unconsciously perceived information for brief durations – on the order of hundreds of milliseconds – is largely uncontroversial (but see Phillips & Block, 2016). Here, we follow most experimentalists working on consciousness by defining visual conscious perception as the subjective experience or visibility of stimuli. Perceptual consciousness can be operationalized (measured) either by objective or subjective measures (Seth, Dienes, Cleeremans, Overgaard, & Pessoa, 2008).

Researchers have used a variety of experimental paradigms to demonstrate unconscious perception (Kim & Blake, 2005). In a standard masked-priming experiment, for example, stimuli are presented briefly and masked so that they are rendered invisible; yet such stimuli are nonetheless thought to be perceived unconsciously because they *prime* or affect downstream behavioral responses (Kouider & Dehaene, 2007). In some experimental paradigms, stimuli are masked and presented for longer durations (Persuh, Emmanouil, & Ro, 2016; Tsuchiya & Koch, 2005). Some type of memory store is implicated in such studies, as behavioral responses are performed in the absence of the perceived objects. We have strong experimental evidence for unconscious response inhibition, a form of cognitive control (van Gaal, de Lange, & Cohen, 2012; van Gaal, Ridderinkhof, Fahrenfort, Scholte, & Lamme, 2008; van Gaal, Ridderinkhof, Scholte, & Lamme, 2010) and for unconsciously deployed metacognitive judgments (Charles, Van Opstal, Marti, & Dehaene, 2013). These are higher order cognitive functions, closely associated with WM. Recently, direct evidence for unconscious iconic memory storage has been provided (Sergent et al., 2013; Thibault, van den Berg, Cavanagh, & Sergent, 2016; Xia, Morimoto, & Noguchi, 2016). Why, then, do so many assume a link between consciousness and VWM?

3.1. Associating Consciousness and VWM

Perhaps the central reason for this assumption is that many maintain that there is a commonsense tie between WM and consciousness. Stein, Kaiser, and Hessemann (2016), for example, write that:

> WM corresponds well to our everyday phenomenology of 'keeping in mind' some information over a short period of time. From this phenomenology, it seems clear that WM is intricately interwoven with conscious awareness. It is difficult to imagine a situation in which we are not consciously aware of the stimuli that enter WM (p. 1).

That many assume a folk-psychological connection between consciousness and WM is consistent with the long history within consciousness studies of assuming that many or even all high-level mental activity requires consciousness (for review see D. M. Rosenthal, 2008; Shea & Frith, 2016).

Unsurprisingly, then, many models of consciousness or VWM implicitly build one phenomenon into the other. Baddeley (2000), for example, modified his original multicomponent model to include an episodic-buffer, which he conceives of as acting as a global workspace (GWS) in Baars' (1988) terminology. The GWS is purportedly a central neural module, which enjoys long-range connections to many areas of the brain; it is thus capable of making information encoded in it available for widespread impact on many neural functions and behavior. According to GWS theories of consciousness (Baars, 1988; Dehaene & Naccache, 2001), the representations in the GWS determine the contents of consciousness. Much neuroimaging data supports GWS theories, purportedly showing that the difference between conscious and unconscious perception consists in differing activations of frontal/parietal areas and widespread connections to other areas (but see Siclari et al., 2017). Although on Baar's (1997) own view the contents of WM need not necessarily be conscious, on Baddeley's GWS-based model of WM, the contents of VWM are invariably conscious (cf. Carruthers, 2015).

Similarly, Prinz's (2000, 2012) attended intermediate-level representation theory of consciousness identifies the contents of consciousness with appropriately attended representations at the intermediate level inspired by Marr's (1982) and Jackendoff's (1987) models of the visual system. And since Prinz equates the relevant kind of attention with the gateway to WM, he likewise holds that only those representations that are available to WM are conscious. Such accounts thereby rule out the unconscious operation of VWM. There are, however, many theoretical reasons to be skeptical that VWM must take as input conscious perception only.

3.2. Dissociating Consciousness and VWM

First, we note that the functional characterization of VWM offered at the outset – that is, a limited-capacity system that allows one to store and manipulate information on the order of seconds – is theoretically neutral insofar as it does not invoke consciousness. If this characterization of VWM is fair, then it remains an open experimental question whether unconsciously perceived information can be encoded in VWM.

Indeed, it is far from obvious that folk psychology includes a tie between WM and consciousness. Since common sense admits of unconscious perceptual states, it would seem at least open that it also includes the possibility that we can "keep in mind" unconsciously perceived information over a short time. Even if Stein, Kaiser, and Hessemann were correct that it is somewhat difficult to generate clear examples of such events, it may simply be because such encoding in VWM is unconscious – and so we are unaware from the first-person perspective that this encoding is occurring.

The assumption that higher mental functions require consciousness is increasingly suspect (Shea & Frith, 2016). And many independently motivated accounts of consciousness do not involve any assumptions about the nature of the contents available to WM. For example, Tononi's (2004) integrated-information theory, according to which mental states are conscious just in case they reach a suitable level of information integration, does not theoretically require that visual contents encoded in VWM exceed the relevant threshold of integration. Likewise, higher-order theories of consciousness, according to which a mental state is conscious just in case one is suitably aware of oneself as being in it (Armstrong, 1968; D. Rosenthal, 2005), provide no reason to think that we must be so aware of the contents that are made available to VWM. These theories of consciousness have experimental support (Lau & Rosenthal, 2011; Tononi, Boly, Massimini, & Koch, 2016) and are consistent with unconscious VWM.

With the development of state-based models of WM, the link between consciousness and WM can be reformulated by asking whether a specific WM state only takes as input conscious information. It seems likely that nonattended information in activated LTM can be represented unconsciously. It has been suggested that perhaps different states of WM correspond to conscious (FoA) and unconscious (activated LTM) information, rather than attended and unattended information (Silvanto, 2017). It might seem obvious that information in the focus of attention (FoA) should always be conscious, but that is not necessary so, as we can attend to unconscious stimuli (Norman, Heywood, & Kentridge, 2013).

The possibility of unconscious VWM is thus interesting in at least two ways. First, it remains an independently interesting question to determine what, if any, mental functions must occur consciously (Berger, 2014). Second, convincing evidence of unconscious WM would require reevaluation and perhaps even rejection of some theories of consciousness.

4. Current Experimental Evidence for Unconscious VWM

Several studies have attempted to demonstrate unconscious VWM.

4.1 Unconscious Operation of VWM

In perhaps the earliest explicit attempt to demonstrate the unconscious operation of VWM, Hassin and colleagues (Hassin, Bargh, Engell, & McCulloch, 2009) presented participants with a rapid series of disks and participants were required to press a button to indicate whether the disks were filled or unfilled. Unbeknownst to the participants, in some conditions the series of disks formed a pattern, which would indicate whether or not a forthcoming disk would be filled. Although participants in the pattern conditions were not able to report on these patterns, they were faster at determining whether the disks were filled than in the non-pattern conditions.

Hassin and colleagues argue that this task required the unconscious operation of VWM because it was necessary for participants to hold in memory a series of disks and compare them to visible disks to determine whether or not they formed a pattern, even though participants did not consciously hold or compare the disks in memory.

Proponents of the view that the contents of VWM are necessarily conscious have offered various critiques of this study. Prinz (2012, p. 96), for example, proposes reasonably that since the stimuli are quite complex, they likely outstrip the limited capacity of VWM and instead implicate fragile VSTM (cf. Carruthers, 2015, p. 86).

But whether or not these stimuli engage only VSTM, it is crucial to note that in all five of these experiments the stimuli were fully visible. Thus, as Carruthers (2015, p. 86) observes, such results arguably demonstrate only that the computed pattern and the resultant expectations are not among the contents of WM, and that this does not show that VWM encodes unconsciously perceived information. And similar considerations likewise undermine what may seem to be evidence of unconscious VWM from experiments involving implicit change detection. Several studies have reported unconscious change detection: for example, studies have found that implicit change detection in the orientation of an item influenced performance on subsequent orientation-judgment tasks (Fernandez-Duque & Thornton, 2000). At first sight, this might seem like *de facto* evidence of unconscious VWM. However, stimuli in these tasks were also fully visible and the delay between two displays was only 250 ms, which arguably is an interval that taps into other types of VSTM rather than VWM.

In other words, reflection on this work reveals an important distinction in types of study of unconscious WM. First, there is the question of whether or not the manipulations of VWM content require consciousness – what we call henceforth the 'unconscious operation' of VWM. This is, however, distinct from the question of whether or not unconsciously perceived information can be encoded in VWM. This latter question is our central interest here.

Other recent studies have likewise provided evidence for the unconscious *operation* of VWM, albeit more indirectly (Bona, Cattaneo, Vecchi, Soto, & Silvanto, 2013; Bona & Silvanto, 2014). Bona and colleagues (2013), for example, examined the relationship between performance and conscious experience in VSTM task. A memory cue was followed, after a delay, by a probe stimulus and participants reported the orientation of probe relative to the memory cue in a forced-choice procedure. After performing a discrimination task, participants reported their conscious experience of the cue stimulus. On half of the trials, masked distractors were presented during the delay period and participants also rated their conscious experience of distractors. Data from this study revealed a double dissociation between performance and conscious experience. Discrimination performance was negatively affected only when distractors differed significantly from cue orientation, regardless of distractor visibility. Cue visibility showed the opposite pattern: visibility was unaffected by distractor orientation. Cue visibility ratings were, however, lower for invisible distractors. These results led authors to conclude that the VSTM memory trace, on which performance is based, is different from the content of conscious experience of VSTM. Furthermore, this evidence led to the proposal of separate representation for conscious experience, a so-called 'conscious copy' model of WM introspection (Jacobs & Silvanto, 2015).

Although cue stimuli in VSTM task in Bona and colleagues (2013) were fully visible and thus cannot directly reveal whether or not unconscious content of VWM is possible; a double dissociation, if independently confirmed, would nonetheless provide indirect evidence for the unconscious operation of VWM. But a simpler explanation of their results would be as follows.

The distractor images modified the cue memory representations – with larger deviations in distractor orientations causing larger shifts in cue memory representations – which in turn affected forced-choice discrimination performance while leaving the vividness of the conscious experience of cue memory intact. In other words, one's conscious experience of a cue is different for different distractor orientations, but equally vivid; one can represent different orientations, equally vividly. It is also possible that invisible distractors modify the vividness of cue memory without affecting forced-choice discrimination performance. Even with lower visibility, participants might have enough orientation information to sustain performance. This reasoning would explain double dissociation results without invoking additional "conscious copy" representation in WM.

We turn in the next section to studies that more directly assess the question of whether or not unconsciously perceived information can make it into VWM.

4.2 Unconscious Content of VWM

In a series of four experiments, Soto, Mäntylä, and Silvanto (2011) presented participants with a masked Gabor patch, to prevent the patch from being consciously perceived. Participants were instructed to keep this cue in memory. After a delay of several seconds, a second Gabor patch, the target, was presented. Participants were then asked to perform cue-target orientation discrimination and to report their awareness of the cue on a scale from 1-4, with 1 indicating no visibility. In some experiments, a distractor was presented after the cue or participants were presented with two cues. In all four experiments, orientation discrimination was above chance level (50%) for trials with visibility ratings of 1.

The authors suggested that data support the existence of unconscious VWM rather than a mere priming effect mainly due to the above-chance performance despite the presence of a distractor and because the gap between cue and target was five seconds in some experiments. These factors purport to show that the cues were held in memory during an ongoing task, and so in VWM.

Some critics of unconscious VWM allege that these results can be explained without appeal to it. Prinz (2012, p. 86), for example, urges that the fact that Soto and colleagues did not find any decrease in performance even after delays of up to five seconds suggests that fragile VSTM, and not VWM, is implicated insofar as VWM putatively shows signs of decay at around four seconds (Zhang & Luck, 2009). By contrast, Carruthers (2015, pg. 87) proposes that deploying attention to unconsciously perceived stimuli might increase the processing of that signal without requiring that the stimuli be encoded in VWM. That is, one might urge, *contra* the authors, that even though participants held the cue in mind over a delay and performed a distractor task, since the information was not in any way manipulated, participants' increased performance was a mere priming effect, requiring storage only in VSTM. Unlike STM, which involves only storage, WM involves not only the maintenance, but also the potential *manipulation* of information (Baddeley, 1986; Luck & Vogel, 2013).

As Stein, Kaiser, and Hesselmann (2016, p. 2) observe, however, a more pressing problem for these experiments is that they depended upon verbal reports of cue visibility. Although such subjective measures are often thought to better reflect perceptual consciousness than objective ones such as forced-choice discrimination (Dehaene & Changeux, 2011; Merikle, Smilek, & Eastwood, 2001), it is well known that such reports are prone to response bias (Peters, Ro, & Lau, 2016; Schmidt, 2015). In other words, trials rated 1 (no visibility) might reflect weak conscious perception of cues, which participants simply fail to report because of conservative standards for regarding a stimulus as seen. Supplementary data demonstrate that participants reported conscious perception of cue on roughly half of the trials. Thus supposing, for a hypothetical experiment, that participants reported conscious perception on large majority (e.g., > 90%) of trials, only 10% of trials would be analyzed. Clearly such evidence of unconscious VWM would be met with skepticism.

For that reason, objective measures of perception, such as d' (the signal-to-noise ratio; MacMillan & Creelman, 2005), are typically preferred. Although Soto and colleagues (2011, p. R912) reported d', they based their calculation only on trials with rating of 1 and thus, as noted by Stein, Kaiser, and Hesselmann (2016, p. 2), their reported pseudo-d' is not a bias-free measure. Although in their reply to Stein and colleagues, Soto and Silvanto (2016) report the actual d', it is not clear whether it is statistically significant.

It is also important to note that most studies of unconscious VWM implicitly assume a slot model and that, according to continuous-resource models, the fundamental metric is not the number of objects stored but instead a precision of each representation, which can vary between items and between trials. It is thus more plausible to suggest that low (although above chance) performance in Soto and colleagues' study stems from the noisy encoding of cues, which were presented briefly and then masked.

Although not discussed by Stein and colleagues, more recent studies (Bergstrom & Eriksson, 2014, 2015; Dutta, Shah, Silvanto, & Soto, 2014; King, Pescetelli, & Dehaene, 2016; Trubutschek et al., 2017) have employed similar approaches and thus suffer from the same concerns about subjective measures of stimulus visibility. Trubutschek and colleagues (2017), for example, used a masking paradigm to render stimuli invisible in a spatial delayed-response task and collected behavioral as well as magnetoencephalography (MEG) data in perception and WM paradigms. Their study set out to address two major concerns facing previous studies of unconscious WM: (1) Participants in the previous studies could have erroneously reported weakly perceived targets as unseen (the 'miscategorization hypothesis') and (2) participants could have made immediate guesses about the target and maintain these guesses in conscious WM (the 'conscious maintenance hypothesis'). Both hypotheses suggest that the results of previous studies could be due to conscious WM. To test these hypotheses, Trubutschek and colleagues examined event-related fields, performed time-frequency analyses, and used machine-learning approaches to dissect neural activity on seen and unseen trials. Importantly, if the miscategorization or conscious maintenance hypotheses were correct, neural signatures on unseen correct trials would resemble neural signatures on seen trials. The location of subjectively unseen targets was reported above chance, seemingly confirming earlier reports of unconscious WM (Soto et al., 2011). Both conscious perception and conscious WM showed shared brain signatures; classifiers trained to separate unseen and seen trials were able to generalize from one task to the other. Furthermore, conscious perception and conscious WM were characterized by sustained desynchronization in the alpha/beta band over frontal cortex and a decodable representation of target location in posterior cortex. Importantly, such activity was not demonstrated for targets on subjectively unseen trials and classifier generalization was unsuccessful. These results provide evidence for unconscious WM, possibly suggesting that synaptic mechanisms support unconscious WM (Mongillo et al., 2008).

One recent model of unconscious WM supported by synaptic mechanisms suggests that WM does not implicate attention, but instead distinct states of WM possibly representing conscious and unconscious information (Silvanto, 2017). According to the model, retro-cues

(Oberauer, 2001) or TMS pulse (Rose et al., 2016) may bring non-cued WM content to conscious experience. This model is consistent with findings showing that we can attend to unconscious information (Norman et al., 2013) and that items in unconscious WM are resistant to distractor interference, which requires attention. This is certainly an interesting possibility that awaits confirmation.

Trubutschek and colleagues note that one of the criteria for WM is manipulation of stored information (Baddeley, 1986; Luck & Vogel, 2013) and that the content of putative unconscious WM was not manipulated in their study. However, the major problem with their study is that, as in previous studies of unconscious WM, consciousness was measured using subjective reports. Although MEG evidence showed desynchronization in the alpha/beta band only for seen trials and not for unseen correct trials, the masking procedure can create a variety of visual experiences that do not necessarily map onto response options. For example, for briefly presented targets in studies using metacontrast masking, a target might change the appearance (e.g., brightness) of the mask, without being perceived (Bachmann & Francis, 2014). In such a case, participants would have some location information in the absence of the conscious perception of the target, supporting above-chance performance on subjectively unseen trials without unconscious WM. If this possibility could be ruled out, we would have a strong evidence for unconscious WM.

Some studies of unconscious WM have, however, not relied on subjective measures of stimuli invisibility. Using the method of the breaking of continuous-flash suppression (CFS), Pan, Lin, Zhao and Soto (2014) reported biasing of visual perception by cues held in unconscious VWM. In CFS, a rapidly changing pattern of Mondrian patterns presented to one eye suppresses conscious perception of stimuli presented to the other eye for several seconds (Tsuchiya & Koch, 2005). Pan and colleagues instructed participants to hold in memory a face cue, which was rendered unconscious by a pattern mask. Signal-detection analysis showed that d' was not significantly different from zero. Using CFS to suppress target face processing, the contrast of the target face in the suppressed eye was gradually increased until participants consciously perceived the face and reported its location. Interestingly, when the target face matched the initial face cue held in VWM, the participants' reaction times were quicker. In a series of control experiments, the authors showed that this effect occurs only when the memory cue is maintained in VWM.

Although Pan and colleagues (2014) did report that the objective measure of consciousness, d', showed that stimuli were not consciously perceived, their experiments did not require participants to manipulate the remembered information in any way. In previous studies (Soto et al., 2011), participants compared the orientation of a stimulus putatively held in unconscious VWM to a target stimulus. In this study, by contrast, unconscious information simply influenced visual processing. Although some type of memory was clearly involved, this experiment is thereby subject to the kind of criticism that Prinz leveled at Hassin and colleagues' (2009) study: that the storage does not clearly meet the minimal requirements of VWM, which involve not only the maintenance, but also the manipulation of information.

Recently, Bergstrom and Eriksson (2017), conducted an fMRI study and used objective measure to assess participants' awareness of memory items suppressed with CFS. Participants performed delayed match-to-sample task in three conditions: a baseline condition with CFS mask only, a conscious condition with objects only, and an unconscious condition in which objects were suppressed with CFS. Participants were first tested in the pre-fMRI session with a 5 s delay period and then in an fMRI session with a 5-15 s delay period. On each trial, participants first performed recognition task, followed by YES/NO detection response, and finally they rated their visual experience on perceptual awareness scale. Only trials with a rating of 1 (no perceptual experience) on the scale were selected for the analysis of the unconscious condition. Although memory performance (d') on the recognition and detection tasks was above chance during the pre-fMRI session, neither was better than chance during the fMRI session. Multivariate pattern analyses of fMRI data from unconscious condition could classify presence versus absence of memory items in prefrontal and occipital cortex, demonstrating the maintenance of unconsciously presented memory items. The authors further suggested that maintenance of unconscious representations in their study depended on persistent neural activity, above activity-silent synaptic changes, and the results are therefore inconsistent with the model of unconscious WM proposed by Trubutschek and colleagues (2017).

One difficulty with this study is that we have no behavioral evidence for the maintenance and manipulation of information, which figures in the operational definition of WM. Although the authors used objective measures of awareness and showed that behavioral performance (d') was at chance, the delay period during the fMRI session was long (5-15 s), and likely contributed to decrease in performance; the authors acknowledged that this and other factors related to fatigue could have played a role in explaining performance. In future experiments, the authors could chose to include a set of randomly intermixed trials on which performance is assessed immediately after the presentation of stimuli.

A set of studies by Rosenthal and colleagues (C. R. Rosenthal, Andrews, Antoniades, Kennard, & Soto, 2016; C. R. Rosenthal, Kennard, & Soto, 2010; C. R. Rosenthal & Soto, 2016) on learning higher-order visuospatial sequences, in the absence of perceptual awareness, provides additional evidence relevant for understanding the relationship between consciousness and WM. The authors used a dichoptic masking protocol, to prevent conscious perception of a complex second-order visuospatial sequence, which was presented repeatedly during the learning phase of the experiment across four monocular locations. Participants were then prompted to discriminate old from new sequences and to rate their confidence during the recognition phase of experiment that followed 20 minutes later. A control experiment revealed that participants were at chance in reporting the eye of origin for individual sequence stimuli. Although participants were at chance at discriminating old from new sequences, confidence ratings revealed that learning did occur. Recognition memory was associated with V1 activity, as a part of network that included the hippocampus. Because learning of visuospatial sequences requires maintenance and manipulation of information over several seconds, thereby meeting the operational definition of WM, these results provide evidence for unconscious WM. Furthermore, because recognition memory, that was associated with hippocampal activation, was probed after a significant delay, these results also reinforce a strong connection between WM and LTM systems.

Two issues related to these findings are worth mentioning. First, visuospatial sequences were repeatedly presented to participants and tested for recognition 20 minutes later. This methodology differs from typical WM tasks in which target stimuli are presented only once and tested for recognition several seconds later. Repeated presentation of stimuli arguably can induce learning through mechanisms independent of WM. The second issue concerns the measurement of recognition memory. Here, the authors showed that both accuracy and sensitivity (d') analyses showed no evidence of learning. It was the usage of metacognitive measures, confidence ratings and type-2 sensitivity, that demonstrated learning. The relationship between perceptual consciousness and metacognitive awareness is not well understood; however, some recent studies suggest that metacognitive judgement is possible in the absence of perceptual sensitivity or awareness (Charles et al., 2013; Jachs, Blanco, Grantham-Hill, & Soto, 2015; Scott, Dienes, Barrett, Bor, & Seth, 2014).

A recent study by Samaha and colleagues (Samaha, Barrett, Sheldon, LaRocque, & Postle, 2016) examined the relationship between metacognitive awareness and WM. Participants performed perceptual and WM tasks for stimuli that were matched for performance (d'), but that varied in confidence ratings (metacognitive awareness). If WM depends on consciousness, then on trials with higher confidence ratings WM performance should improve. This hypothesis naturally depends on assumption that metacognitive awareness is a good measure of perceptual consciousness. But the authors found no evidence for this hypothesis, suggesting that WM is independent of conscious perception. Although these results are suggestive, it should be noted that one does need an additional theoretical assumption to conclude that WM can store unconsciously perceived items. Some minimal perceptual consciousness of to-be-remembered items might be required for WM, even if further increase in metacognitive awareness does not improve WM performance.

An important issue raised by studies that employ metacognitive judgements concerns the relationship between perceptual consciousness and metacognitive awareness. Jachs and colleagues (2015) have demonstrated that whereas stimulus awareness has a strong effect on perceptual discrimination (d'), the effect on confidence judgements (metacognition) was much weaker. These results suggest that perceptual and metacognitive judgements do not operate on the same input and that metacognition is not tightly linked to perceptual consciousness. Converging evidence comes from studies on error detection. Charles and colleagues (2013) showed that participants were able to detect their errors above chance even under subliminal conditions. They proposed that two distinct mechanisms exist for metacognitive judgements; conscious and unconscious evaluation systems. Corroborating evidence from a recent study that evaluated error-monitoring performance in schizophrenia patients and healthy controls showed that only conscious metacognition is affected in schizophrenia, whereas unconscious monitoring performance remained intact (Charles et al., 2017). These results are important because they suggest that (1) if metacognition can be dissociated from perceptual awareness, then we cannot use metacognitive judgements to assess perceptual consciousness and (2) since metacognition is considered a higher-order process, it strengthens the evidence that higher-order processes related to WM can be deployed unconsciously.

5. Ways Forward in the Study of Unconscious VWM

Having examined the current literature regarding the relationship between consciousness and VWM, we find that there is no definitive evidence for unconscious VWM. But we see no in principle barrier to the demonstration of unconscious VWM. Thus in closing we offer recommendations for how to move forward in the study of unconscious VWM, in light of the lessons gleaned from the work that has already been done.

In short, because of the problem with response bias, proper experimental design should use forced-choice discrimination as a measure of consciousness and an indirect measure to demonstrate storage and manipulation of information.

Moreover, in order to study the unconscious content of VWM, and not merely VWM's unconscious operation, target stimuli must of course somehow be invisible to visual consciousness. But one well-known pitfall for studies that involve the technique of visual masking is that it typically involves brief presentation times and masks that inevitably degrade the stimuli, thereby reducing their signal strength. Consequently, some theorists have speculated that the fact that we have not decisively experimentally demonstrated many higher cognitive functions occurring unconsciously may be an artifact of our current methods for masking stimuli (e.g., Lau, 2009; Persuh et al., 2016). Thus, were a legitimate study of unconscious VWM involving masking to be devised, the failure of participants to successfully perform the memory task might be explained not by the fact that the unconsciously information cannot be encoded in VWM, but by the fact that unconscious perception, as it is currently studied, is typically weak. New experimental techniques might be developed to explore the full extent of unconscious perception and unconscious VWM (Persuh et al., 2016; Tsuchiya & Koch, 2005).

For these reasons, blindsight – wherein people with cortical damage to V1 are capable of discriminating stimuli in their blind regions (scotoma) though they report no perceptual consciousness (Weiskrantz, 1986) – may seem to provide a particularly promising route for studying unconscious VWM. In blindsight patients, high contrast stimuli can be presented for unlimited durations. Skeptics regarding unconscious VWM might regard blindsight responses as drawing on a type of fast and automatic processing akin to priming, but most participants in studies of blindsight are not required to make speeded responses; instead they have several seconds to respond to visibility and forced-choice discrimination questions. To respond correctly, unconscious visual information must be stored and VWM seems a likely candidate. To our knowledge, however, no study of blindsight has utilized a standard task for testing VWM, such as requiring participants to manipulate unconsciously perceived information after a sizable delay period.

We recognize, however, that there are several criticisms of blindsight, such as the fact that it arguably involves a form of degraded normal vision (Weiskrantz, 2009) or that a participant might unconsciously perceive the stimulus, on that basis make a conscious guess, and then retain the conscious representation in memory before responding (Stein et al., 2016). These criticisms can certainly be addressed with proper experimental design. Although blindsight is a paradigmatic example of a dissociation between objective and subjective measures of consciousness, a modified experimental design could employ a combination of objective measure of consciousness and an indirect measure of storage and manipulation of information.

One might, however, prefer evidence of unconscious VWM in healthy individuals. Lau and colleagues coined the expression ‘relative blindsight’ to describe differing levels of subjective visibility for stimuli in healthy individuals that were nonetheless matched for objective task performance (Lau & Passingham, 2006). It has been proposed that this approach could be a fruitful way to study unconscious WM (Samaha, 2015). Although the original study faced several criticisms (Balsdon & Azzopardi, 2015; Jannati & Di Lollo, 2012), which were addressed in subsequent studies (Maniscalco, Peters, & Lau, 2016; Peters et al., 2016), the proponents of the relative blindsight paradigm report that they have abandoned this strategy due to very small, although reproducible effects (Peters et al., 2016). Relative blindsight might thus not be a particularly fruitful avenue for future studies of unconscious VWM. A related concern is that relative blindsight cannot provide direct evidence for unconscious WM, because we cannot exclude the possibility that some level of perceptual consciousness is necessary for WM, even if WM performance is not modulated by further changes in perceptual consciousness.

Some studies with healthy participants employing CFS, by contrast, may have provided indirect evidence for unconscious VWM. Sklar and colleagues (2012), for example, presented participants with an equation masked by CSF for up to two seconds and then asked to verbalize a visible number. When the result of the equation and the number were congruent participants were faster. Individual participants that performed above chance based on binomial distribution analysis were excluded from the analyses. Since recent evidence suggests that VWM is engaged not only for maintenance of visual information but also for items still in view (Tsubomi, Fukuda, Watanabe, & Vogel, 2013), these results arguably suggest the existence of unconscious VWM. There is, however, some concern about the replicability of this study (Karpinski, Yale, & Briggs, 2017). A more definitive study of unconscious VWM in healthy individuals might use a similar CFS-based set-up but use a more traditional task regarding VWM.

Perhaps more pressingly, given the present lack of consensus regarding the nature of WM and its neural basis, it would be reasonable for future studies of unconscious VWM to use set sizes and delay intervals between the presentation of a to-be-remembered stimulus and the memory task that clearly separates VWM from other memory systems such as fragile VSTM.

Lastly, we take it that perhaps the *key* feature of WM, which distinguishes it from STM and LTM, is that information encoded in WM is available for use in ongoing tasks. That is, such information must be available for *manipulation.* To be clear, we do not claim that information stored within WM must be manipulated; it is consistent with current models that such information may simply be stored and forgotten. But, to count as being encoded in VWM, information must at least be disposed to be manipulated. This is why standard paradigms for studying conscious VWM, such as the N-back task or the operation-span task (Conway et al., 2005), often involve consciously holding a to-be-remembered target in mind and comparing it to consciously presented subsequent stimuli. Many purported studies of unconscious VWM that do not involve such manipulation are open to the criticism that the information gleaned from unconsciously perceived targets merely primes participants or is stored in (fragile) VSTM (Soto et al., 2011). Perhaps the manipulation of target information might involve participants' being primed after mentally rotating remembered stimuli (Hyun & Luck, 2007) or after performing arithmetic operations on remembered numbers (cf. Sklar et al., 2012).

Whether or not there are ways to modify standard paradigms for studying WM, any required unconscious manipulation of target information will doubtless increase the difficulty of such tasks, thus reducing the likelihood that participants can successfully perform them. Coupling this with the problem that unconscious stimuli are often weakly encoded, the possibility that unconsciously perceived information could survive the relevant delay period to be manipulated may seem remote. But so far as we see it, there is nothing theoretically that would rule out the possibility of unconscious VWM.

Acknowledgments: We thank Hakwan Lau, Myrto Mylopoulos, Chandra Sripada, Genevieve Feldmann and Kandi Turley-Ames for their helpful comments on earlier drafts of this manuscript.

Funding: This work was supported by a pilot grant from the Center for Consciousness Science, University of Michigan Medical School.

References

Armstrong, D. M. (1968). *A Materialist Theory of the Mind*: Routledge & K. Paul.

Atkinson, R. C., & Shiffrin, R. M. (1968). Human memory: A proposed system and its control processes. In K. W. Spence (Ed.), *The psychology of learning and motivation: Advances in research and theory* (Vol. 2, pp. 89-195). New York: Academic Press.

Awh, E., & Jonides, J. (2001). Overlapping mechanisms of attention and spatial working memory. *Trends Cogn Sci, 5*(3), 119-126.

Baars, B. J. (1988). *A Cognitive Theory of Consciousness*. New York: Cambridge University Press.

Baars, B. J. (1997). Some essential differences between consciousness and attention, perception, and working memory. *Conscious Cogn, 6*(2-3), 363-371. doi: 10.1006/ccog.1997.0307

Bachmann, T., & Francis, G. (2014). *Visual Masking: Studying Perception, Attention, and Consciousness* (1th ed.): Elsevier.

Baddeley, A. D. (1986). *Working Memory*. Clarendon, Oxford: Oxford University Press.

Baddeley, A. D. (2000). The episodic buffer: a new component of working memory? *Trends Cogn Sci, 4*(11), 417-423.

Baddeley, A. D., & Hitch, G. J. (1974). Working memory. In G. H. Bower (Ed.), *The psychology of learning and motivation* (Vol. VIII, pp. 47-90). New York: Academic Press.

Balsdon, T., & Azzopardi, P. (2015). Absolute and relative blindsight. *Conscious Cogn, 32*, 79-91. doi: 10.1016/j.concog.2014.09.010

Berger, J. (2014). Mental states, conscious and nonconscious. *Philosophy Compass, 9*(6), 392-401.

Bergstrom, F., & Eriksson, J. (2014). Maintenance of non-consciously presented information engages the prefrontal cortex. *Front Hum Neurosci, 8*, 938. doi: 10.3389/fnhum.2014.00938

Bergstrom, F., & Eriksson, J. (2015). The conjunction of non-consciously perceived object identity and spatial position can be retained during a visual short-term memory task. *Front Psychol, 6*, 1470. doi: 10.3389/fpsyg.2015.01470

Bergstrom, F., & Eriksson, J. (2017). Neural Evidence for Non-conscious Working Memory. *Cereb Cortex*, 1-12. doi: 10.1093/cercor/bhx193

Bona, S., Cattaneo, Z., Vecchi, T., Soto, D., & Silvanto, J. (2013). Metacognition of Visual Short-Term Memory: Dissociation between Objective and Subjective Components of VSTM. *Front Psychol, 4*, 62. doi: 10.3389/fpsyg.2013.00062

Bona, S., & Silvanto, J. (2014). Accuracy and confidence of visual short-term memory do not go hand-in-hand: behavioral and neural dissociations. *PloS One, 9*(3), e90808. doi: 10.1371/journal.pone.0090808

Carruthers, P. (2015). *The Centered Mind: What the Science of Working Memory Shows Us About the Nature of Human Thought* Oxford University Press.

Charles, L., Gaillard, R., Amado, I., Krebs, M. O., Bendjemaa, N., & Dehaene, S. (2017). Conscious and unconscious performance monitoring: Evidence from patients with schizophrenia. *NeuroImage, 144*(Pt A), 153-163. doi: 10.1016/j.neuroimage.2016.09.056

Charles, L., Van Opstal, F., Marti, S., & Dehaene, S. (2013). Distinct brain mechanisms for conscious versus subliminal error detection. *NeuroImage, 73*, 80-94. doi: 10.1016/j.neuroimage.2013.01.054

Conway, A. R., Kane, M. J., Bunting, M. F., Hambrick, D. Z., Wilhelm, O., & Engle, R. W. (2005). Working memory span tasks: A methodological review and user's guide. *Psychon Bull Rev, 12*(5), 769-786.

Courtney, S. M., Ungerleider, L. G., Keil, K., & Haxby, J. V. (1997). Transient and sustained activity in a distributed neural system for human working memory. *Nature, 386*(6625), 608-611. doi: 10.1038/386608a0

Cowan, N. (1995). Attention and Memory: An Integrated Framework. New York: Oxford University Press.

Cowan, N. (2001). The magical number 4 in short-term memory: a reconsideration of mental storage capacity. *The Behavioral and Brain Sciences, 24*(1), 87-114; discussion 114-185-187-114; discussion 114-185.

D'Esposito, M., & Postle, B. R. (2015). The cognitive neuroscience of working memory. *Annu Rev Psychol, 66*, 115-142. doi: 10.1146/annurev-psych-010814-015031

Dehaene, S., & Changeux, J.-P. (2011). Experimental and theoretical approaches to conscious processing. *Neuron, 70*(2), 200-227. doi: 10.1016/j.neuron.2011.03.018

Dehaene, S., & Naccache, L. (2001). Towards a cognitive neuroscience of consciousness: basic evidence and a workspace framework. *Cognition, 79*(1-2), 1-37.

Dutta, A., Shah, K., Silvanto, J., & Soto, D. (2014). Neural basis of non-conscious visual working memory. *NeuroImage, 91*, 336-343. doi: 10.1016/j.neuroimage.2014.01.016

Fernandez-Duque, D., & Thornton, I. M. (2000). Change Detection Without Awareness: Do Explicit Reports Underestimate the Representation of Change in the Visual System? *VISUAL COGNITION, 7*(1-3), 323-344. doi: 10.1080/135062800394838

Fougnie, D., Suchow, J. W., & Alvarez, G. A. (2012). Variability in the quality of visual working memory. *Nat Commun, 3*, 1229. doi: 10.1038/ncomms2237

Fukuda, K., Vogel, E., Mayr, U., & Awh, E. (2010). Quantity, not quality: the relationship between fluid intelligence and working memory capacity. *Psychon Bull Rev, 17*(5), 673-679. doi: 10.3758/17.5.673

Fuster, J. M., & Alexander, G. E. (1971). Neuron activity related to short-term memory. *Science, 173*(3997), 652-654.

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

Hassin, R. R., Bargh, J. A., Engell, A. D., & McCulloch, K. C. (2009). Implicit working memory. *Conscious Cogn, 18*(3), 665-678. doi: 10.1016/j.concog.2009.04.003

Hyun, J. S., & Luck, S. J. (2007). Visual working memory as the substrate for mental rotation. *Psychon Bull Rev, 14*(1), 154-158.

Jachs, B., Blanco, M. J., Grantham-Hill, S., & Soto, D. (2015). On the independence of visual awareness and metacognition: a signal detection theoretic analysis. *J Exp Psychol Hum Percept Perform, 41*(2), 269-276. doi: 10.1037/xhp0000026

Jackendoff, R. (1987). *Consciousness and the Computational Mind*. Cambridge, MA: MIT Press.

Jacobs, C., & Silvanto, J. (2015). How is working memory content consciously experienced? The 'conscious copy' model of WM introspection. *Neurosci Biobehav Rev, 55*, 510-519. doi: 10.1016/j.neubiorev.2015.06.003

Jannati, A., & Di Lollo, V. (2012). Relative blindsight arises from a criterion confound in metacontrast masking: implications for theories of consciousness. *Conscious Cogn, 21*(1), 307-314. doi: 10.1016/j.concog.2011.10.003

Johnson, M. K., McMahon, R. P., Robinson, B. M., Harvey, A. N., Hahn, B., Leonard, C. J., . . . Gold, J. M. (2013). The relationship between working memory capacity and broad measures of cognitive ability in healthy adults and people with schizophrenia. *Neuropsychology, 27*(2), 220-229. doi: 10.1037/a0032060

Karpinski, A., Yale, M., & Briggs, J. C. (2017). Retraction statement: Unconscious arithmetic processing: A direct replication. *European Journal of Social Psychology, 47*(1), 114.

Kim, C.-Y., & Blake, R. (2005). Psychophysical magic: rendering the visible 'invisible'. *Trends In Cognitive Sciences, 9*(8), 381-388. doi: 10.1016/j.tics.2005.06.012

King, J. R., Pescetelli, N., & Dehaene, S. (2016). Brain Mechanisms Underlying the Brief Maintenance of Seen and Unseen Sensory Information. *Neuron, 92*(5), 1122-1134. doi: 10.1016/j.neuron.2016.10.051

Kouider, S., & Dehaene, S. (2007). Levels of processing during non-conscious perception: a critical review of visual masking. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences, 362*(1481), 857-875.

Kubota, K., & Niki, H. (1971). Prefrontal cortical unit activity and delayed alternation performance in monkeys. *J Neurophysiol, 34*(3), 337-347.

Larocque, J. J., Lewis-Peacock, J. A., & Postle, B. R. (2014). Multiple neural states of representation in short-term memory? It's a matter of attention. *Front Hum Neurosci, 8*, 5. doi: 10.3389/fnhum.2014.00005

Lau, H. C. (2009). Volition and the function of consciousness. In N. Murphy, G. F. R. Ellis & T. O'Connor (Eds.), *Downward Causation and the Neurobiology of Free Will*. Berlin: Springer.

Lau, H. C., & Passingham, R. E. (2006). Relative blindsight in normal observers and the neural correlate of visual consciousness. *Proceedings Of The National Academy Of Sciences Of The United States Of America, 103*(49), 18763-18768. doi: 10.1073/pnas.0607716103

Lau, H. C., & Rosenthal, D. (2011). Empirical support for higher-order theories of conscious awareness. *Trends In Cognitive Sciences, 15*(8), 365-373. doi: 10.1016/j.tics.2011.05.009

Lewis-Peacock, J. A., Drysdale, A. T., Oberauer, K., & Postle, B. R. (2012). Neural evidence for a distinction between short-term memory and the focus of attention. *J Cogn Neurosci, 24*(1), 61-79. doi: 10.1162/jocn_a_00140

Luck, S. J., & Vogel, E. K. (1997). The capacity of visual working memory for features and conjunctions. *Nature, 390*(6657), 279-281.

Luck, S. J., & Vogel, E. K. (2013). Visual working memory capacity: from psychophysics and neurobiology to individual differences. *Trends Cogn Sci, 17*(8), 391-400. doi: 10.1016/j.tics.2013.06.006

Ma, W. J., Husain, M., & Bays, P. M. (2014). Changing concepts of working memory. *Nat Neurosci, 17*(3), 347-356. doi: 10.1038/nn.3655

MacMillan, N. A., & Creelman, C. D. (2005). *Detection Theory: A User's Guide*: Erlbaum.

Magnussen, S. (2000). Low-level memory processes in vision. *Trends Neurosci, 23*(6), 247-251.

Maniscalco, B., Peters, M. A., & Lau, H. (2016). Heuristic use of perceptual evidence leads to dissociation between performance and metacognitive sensitivity. *Atten Percept Psychophys, 78*(3), 923-937. doi: 10.3758/s13414-016-1059-x

Marr, D. (1982). *Vision: A Computational Investigation into the Human Representation and Processing of Visual Information*. Cambridge, MA: MIT Press.

Matsukura, M., & Hollingworth, A. (2011). Does visual short-term memory have a high-capacity stage? *Psychon Bull Rev, 18*(6), 1098-1104. doi: 10.3758/s13423-011-0153-2

McElree, B. (1998). Attended and non-attended states in working memory: accessing categorized structures. *J. Mem. Lang., 38*, 225-252.

Merikle, P. M., Smilek, D., & Eastwood, J. D. (2001). Perception without awareness: perspectives from cognitive psychology. *Cognition, 79*(1-2), 115-134.

Miller, G. A. (1956). The magical number seven, plus or minus two: some limits on our capacity for processing information. *Psychol. Rev., 63*, 81–97.

Miller, G. A., Galanter, E., & Pribram, K. H. (1960). *Plans and the Structure of Behavior*. New York, New York: Holt, Rinehart & Winston.

Mongillo, G., Barak, O., & Tsodyks, M. (2008). Synaptic theory of working memory. *Science, 319*(5869), 1543-1546. doi: 10.1126/science.1150769

Norman, L. J., Heywood, C. A., & Kentridge, R. W. (2013). Object-based attention without awareness. *Psychol Sci, 24*(6), 836-843. doi: 10.1177/0956797612461449

Oberauer, K. (2001). Removing irrelevant information from working memory: a cognitive aging study with the modified Sternberg task. *J Exp Psychol Learn Mem Cogn, 27*(4), 948-957.

Oberauer, K. (2002). Access to information in working memory: exploring the focus of attention. *J Exp Psychol Learn Mem Cogn, 28*(3), 411-421.

Pan, Y., Lin, B., Zhao, Y., & Soto, D. (2014). Working memory biasing of visual perception without awareness. *Atten Percept Psychophys, 76*(7), 2051-2062. doi: 10.3758/s13414-013-0566-2

Persuh, M., Emmanouil, T. A., & Ro, T. (2016). Perceptual overloading reveals illusory contour perception without awareness of the inducers. *Atten Percept Psychophys, 78*(6), 1692-1701. doi: 10.3758/s13414-016-1146-z

Peters, M. A., Ro, T., & Lau, H. (2016). Who's afraid of response bias? *Neurosci Conscious, 2016*(1). doi: 10.1093/nc/niw001

Phillips, I., & Block, N. (2016). Debate on unconscious perception. In B. Nanay (Ed.), *Current Controversies in Philosophy of Perception* (pp. 165–192). New York: Routledge.

Prinz, J. J. (2000). A neurofunctional theory of visual consciousness. *Conscious Cogn, 9*(2 Pt 1), 243-259. doi: 10.1006/ccog.2000.0442

Prinz, J. J. (2012). *The Conscious Brain: How Attention Engenders Experience*: Oxford University Press.

Rose, N. S., LaRocque, J. J., Riggall, A. C., Gosseries, O., Starrett, M. J., Meyering, E. E., & Postle, B. R. (2016). Reactivation of latent working memories with transcranial magnetic stimulation. *Science, 354*(6316), 1136-1139. doi: 10.1126/science.aah7011

Rosenthal, C. R., Andrews, S. K., Antoniades, C. A., Kennard, C., & Soto, D. (2016). Learning and Recognition of a Non-conscious Sequence of Events in Human Primary Visual Cortex. *Curr Biol, 26*(6), 834-841. doi: 10.1016/j.cub.2016.01.040

Rosenthal, C. R., Kennard, C., & Soto, D. (2010). Visuospatial sequence learning without seeing. *PloS One, 5*(7), e11906. doi: 10.1371/journal.pone.0011906

Rosenthal, C. R., & Soto, D. (2016). The Anatomy of Non-conscious Recognition Memory. *Trends Neurosci, 39*(11), 707-711. doi: 10.1016/j.tins.2016.09.005

Rosenthal, D. (2005). *Consciousness and Mind*: Clarendon Press.

Rosenthal, D. M. (2008). Consciousness and its function. *Neuropsychologia, 46*(3), 829-840. doi: 10.1016/j.neuropsychologia.2007.11.012

Samaha, J. (2015). How best to study the function of consciousness? *Front Psychol, 6*, 604. doi: 10.3389/fpsyg.2015.00604

Samaha, J., Barrett, J. J., Sheldon, A. D., LaRocque, J. J., & Postle, B. R. (2016). Dissociating Perceptual Confidence from Discrimination Accuracy Reveals No Influence of Metacognitive Awareness on Working Memory. *Front Psychol, 7*, 851. doi: 10.3389/fpsyg.2016.00851

Schmidt, T. (2015). Invisible Stimuli, Implicit Thresholds: Why Invisibility Judgments Cannot be Interpreted in Isolation. *Adv Cogn Psychol, 11*(2), 31-41. doi: 10.5709/acp-0169-3

Scott, R. B., Dienes, Z., Barrett, A. B., Bor, D., & Seth, A. K. (2014). Blind insight: metacognitive discrimination despite chance task performance. *Psychol Sci, 25*(12), 2199-2208. doi: 10.1177/0956797614553944

Serences, J. T., Ester, E. F., Vogel, E. K., & Awh, E. (2009). Stimulus-specific delay activity in human primary visual cortex. *Psychol Sci, 20*(2), 207-214. doi: 10.1111/j.1467-9280.2009.02276.x

Sergent, C., Wyart, V., Babo-Rebelo, M., Cohen, L., Naccache, L., & Tallon-Baudry, C. (2013). Cueing attention after the stimulus is gone can retrospectively trigger conscious perception. *Current Biology: CB, 23*(2), 150-155. doi: 10.1016/j.cub.2012.11.047

Seth, A. K., Dienes, Z., Cleeremans, A., Overgaard, M., & Pessoa, L. (2008). Measuring consciousness: relating behavioural and neurophysiological approaches. *Trends In Cognitive Sciences, 12*(8), 314-321. doi: 10.1016/j.tics.2008.04.008

Shea, N., & Frith, C. D. (2016). Dual-process theories and consciousness: the case for 'Type Zero' cognition. *Neuroscience of Consciousness, 2016*(1).

Siclari, F., Baird, B., Perogamvros, L., Bernardi, G., LaRocque, J. J., Riedner, B., . . . Tononi, G. (2017). The neural correlates of dreaming. *Nat Neurosci, 20*(6), 872-878. doi: 10.1038/nn.4545

Silvanto, J. (2017). Working Memory Maintenance: Sustained Firing or Synaptic Mechanisms? *Trends Cogn Sci, 21*(3), 152-154. doi: 10.1016/j.tics.2017.01.009

Sklar, A. Y., Levy, N., Goldstein, A., Mandel, R., Maril, A., & Hassin, R. R. (2012). Reading and doing arithmetic nonconsciously. *Proc Natl Acad Sci U S A, 109*(48), 19614-19619. doi: 10.1073/pnas.1211645109

Sligte, I. G., Scholte, H. S., & Lamme, V. A. F. (2008). Are there multiple visual short-term memory stores? *PloS One, 3*(2), e1699-e1699.

Sligte, I. G., Vandenbroucke, A. R. E., Scholte, H. S., & Lamme, V. A. F. (2010). Detailed sensory memory, sloppy working memory. *Frontiers in Psychology, 1*, 175-175.

Soto, D., Mantyla, T., & Silvanto, J. (2011). Working memory without consciousness. *Curr Biol, 21*(22), R912-913. doi: 10.1016/j.cub.2011.09.049

Soto, D., & Silvanto, J. (2016). Is conscious awareness needed for all working memory processes? *Neuroscience of Consciousness, 2016*, 1-3.

Souza, A. S., Rerko, L., Lin, H. Y., & Oberauer, K. (2014). Focused attention improves working memory: implications for flexible-resource and discrete-capacity models. *Atten Percept Psychophys, 76*(7), 2080-2102. doi: 10.3758/s13414-014-0687-2

Sperling, G. (1960). The information available in brief visual presentations. *Psychological monographs: General and applied, 74*(11), 1-29.

Stein, T., Kaiser, D., & Hesselmann, G. (2016). Can working memory be non-conscious? *Neuroscience of Consciousness, 2016*, 1-3.

Thibault, L., van den Berg, R., Cavanagh, P., & Sergent, C. (2016). Retrospective Attention Gates Discrete Conscious Access to Past Sensory Stimuli. *PloS One, 11*(2), e0148504. doi: 10.1371/journal.pone.0148504

Tononi, G. (2004). An information integration theory of consciousness. *BMC Neuroscience, 5*. doi: 10.1186/1471-2202-5-42

Tononi, G., Boly, M., Massimini, M., & Koch, C. (2016). Integrated information theory: from consciousness to its physical substrate. *Nat Rev Neurosci, 17*(7), 450-461. doi: 10.1038/nrn.2016.44

Trubutschek, D., Marti, S., Ojeda, A., King, J. R., Mi, Y., Tsodyks, M., & Dehaene, S. (2017). A theory of working memory without consciousness or sustained activity. *Elife, 6*. doi: 10.7554/eLife.23871

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

Tsuchiya, N., & Koch, C. (2005). Continuous flash suppression reduces negative afterimages. *Nature Neuroscience, 8*(8), 1096-1101. doi: 10.1038/nn1500

van den Berg, R., Shin, H., Chou, W. C., George, R., & Ma, W. J. (2012). Variability in encoding precision accounts for visual short-term memory limitations. *Proc Natl Acad Sci U S A, 109*(22), 8780-8785. doi: 10.1073/pnas.1117465109

van Gaal, S., de Lange, F. P., & Cohen, M. X. (2012). The role of consciousness in cognitive control and decision making. *Frontiers In Human Neuroscience, 6*. doi: 10.3389/fnhum.2012.00121

van Gaal, S., Ridderinkhof, K. R., Fahrenfort, J. J., Scholte, H. S., & Lamme, V. A. F. (2008). Frontal cortex mediates unconsciously triggered inhibitory control. *The Journal Of Neuroscience: The Official Journal Of The Society For Neuroscience, 28*(32), 8053-8062.

van Gaal, S., Ridderinkhof, K. R., Scholte, H. S., & Lamme, V. A. F. (2010). Unconscious activation of the prefrontal no-go network. *The Journal Of Neuroscience: The Official Journal Of The Society For Neuroscience, 30*(11), 4143-4150. doi: 10.1523/jneurosci.2992-09.2010

Weiskrantz, L. (1986). *Blindsight: A Case Study and Implications*. Oxford: Oxford University Press.

Weiskrantz, L. (2009). Is blindsight just degraded normal vision? *Experimental Brain Research. Experimentelle Hirnforschung. Expérimentation Cérébrale, 192*(3), 413-416.

Wolfe, J. M. (2014). Introduction to the special issue on visual working memory. *Atten Percept Psychophys, 76*(7), 1861-1870. doi: 10.3758/s13414-014-0783-3

Xia, Y., Morimoto, Y., & Noguchi, Y. (2016). Retrospective triggering of conscious perception by an interstimulus interaction. *J Vis, 16*(7), 3. doi: 10.1167/16.7.3

Zarahn, E., Aguirre, G., & D'Esposito, M. (1997). A trial-based experimental design for fMRI. *NeuroImage, 6*(2), 122-138. doi: 10.1006/nimg.1997.0279

Zhang, W., & Luck, S. J. (2009). Sudden death and gradual decay in visual working memory. *Psychol Sci, 20*(4), 423-428. doi: 10.1111/j.1467-9280.2009.02322.x