

# Separating memoranda in depth increases visual working memory performance

**Chaipat Chunharas**

Psychology Department,  
University of California San Diego, La Jolla, CA, USA  
King Chulalongkorn Memorial Hospital,  
Chulalongkorn University, Bangkok, Thailand



**Rosanne L. Rademaker**

Psychology Department,  
University of California San Diego, La Jolla, CA, USA  
Donders Institute for Brain, Cognition and Behavior,  
Radboud University, Nijmegen, the Netherlands



**Thomas C. Sprague**

Department of Psychology,  
University of California Santa Barbara,  
Santa Barbara, CA, USA



**Timothy F. Brady**

Psychology Department,  
University of California San Diego, La Jolla, CA, USA



**John T. Serences**

Psychology Department,  
University of California San Diego, La Jolla, CA, USA  
Neurosciences Graduate Program,  
University of California San Diego, La Jolla, CA, USA  
Kavli Institute for Brain and Mind,  
University of California, San Diego, La Jolla, CA, USA



Visual working memory is the mechanism supporting the continued maintenance of information after sensory inputs are removed. Although the capacity of visual working memory is limited, memoranda that are spaced farther apart on a 2-D display are easier to remember, potentially because neural representations are more distinct within retinotopically organized areas of visual cortex during memory encoding, maintenance, or retrieval. The impact on memory of spatial separability in depth is less clear, even though depth information is essential to guiding interactions with objects in the environment. On one account, separating memoranda in depth may facilitate performance if interference between items is reduced. However, depth information must be inferred indirectly from the 2-D retinal image, and less is known about how visual cortex represents depth. Thus, an alternative possibility is that separation in depth does not attenuate between-items interference; it may even impair performance, as attention must be distributed across a larger volume of 3-D space. We tested these alternatives using a stereo display while

participants remembered the colors of stimuli presented either near or far in the 2-D plane or in depth. Increasing separation in-plane and in depth both enhanced performance. Furthermore, participants who were better able to utilize stereo depth cues showed larger benefits when memoranda were separated in depth, particularly for large memory arrays. The observation that spatial separation in the inferred 3-D structure of the environment improves memory performance, as is the case in 2-D environments, suggests that separating memoranda in depth might reduce neural competition by utilizing cortically separable resources.

## Introduction

Visual working memory (VWM) supports the integration of past and present sensory information via short-term maintenance when such information is no longer directly accessible. Performance on VWM tasks

Citation: Chunharas, C., Rademaker, R. L., Sprague, T. C., Brady, T. F., & Serences, J. T. (2019). Separating memoranda in depth increases visual working memory performance. *Journal of Vision*, 19(1):4, 1–16, <https://doi.org/10.1167/19.1.4>.



is highly correlated with measures of general intelligence and other related outcome measures and is therefore thought to reflect a core cognitive capacity (Baddeley, 1986; Conway, Cowan, Bunting, Theriault, & Minkoff, 2002; Engle, Tuholski, Laughlin, & Conway, 1999; Fukuda, Vogel, Mayr, & Awh, 2010). In most VWM studies, simple visual stimuli are presented on a 2-D computer screen and participants remember specific features, such as color or orientation, that are presented at different spatial locations (Engle et al., 1999; Luck & Vogel, 1997; Simons & Levin, 1997; Zhang & Luck, 2008). Based on such work, VWM is known to be limited in capacity (Bays, Catalao, & Husain, 2009; Bays & Husain, 2008; Ma, Husain, & Bays, 2014; Schurgin, Wixted, & Brady, 2018), such that increasing the number of items to be remembered or the delay duration leads to reductions in memory precision (Ma et al., 2014; Panichello, DePasquale, Pillow, & Buschman, 2018; Rademaker, Park, & Sack, 2018; Shin, Zou, & Ma, 2017; van den Berg, Shin, Chou, George, & Ma, 2012; Zhang & Luck, 2008), reductions in confidence (Rademaker, Tredway, & Tong, 2012), the misbinding or “swapping” of different visual features (Bays, 2016; Bays, Gorgoraptis, Wee, Marshall, & Husain, 2011; Bays, Wu, & Husain, 2011), and the tendency to chunk information into group-level ensemble representations (Brady & Alvarez, 2011).

One of the key factors that govern interactions between remembered items is the degree to which different memoranda can be bound to distinct spatial locations. For example, detecting a change in a remembered object is more challenging when the spatial configuration of the display is modified between encoding and test, highlighting the importance of spatial layout and spatial location in VWM (Hollingworth, 2007; Hollingworth & Rasmussen, 2010; Jiang, Olson, & Chun, 2000; Olson & Marshuetz, 2005; Phillips, 1974; Postle, Awh, Serences, Sutterer, & D’Esposito, 2013; Treisman & Zhang, 2006). Memory performance is improved when multiple simultaneous memoranda are presented far from each other, compared to close to each other, suggesting a role for spatial interference (Cohen, Rhee, & Alvarez, 2016; Emrich & Ferber, 2012). Furthermore, presenting memoranda sequentially in different spatial locations leads to better memory performance compared to sequentially presenting items in the same spatial location, even when location is task irrelevant (Pertzov & Husain, 2014).

The importance of 2-D space in VWM is consistent with the clear maplike organization of 2-D spatial position across the cortical surface, which should result in less neural competition and more distinct representations as items are spaced farther apart (Engel, Glover, & Wandell, 1997; Grill-Spector & Malach, 2004; Maunsell & Newsome, 1987; Sereno et al., 1995;

Sereno, Pitzalis, & Martinez, 2001; Talbot & Marshall, 1941). This general idea is consistent with a sensory-recruitment account, which proposes that early sensory cortex supports the maintenance of sensory information in working memory (D’Esposito & Postle, 2015; Emrich, Riggall, Larocque, & Postle, 2013; Harrison & Tong, 2009; Pasternak & Greenlee, 2005; Rademaker, Chunharas, & Serences, 2018; Serences, 2016; Serences, Ester, Vogel, & Awh, 2009; Sreenivasan, Curtis, & D’Esposito, 2014). Thus, overlap or competition between representations in retinotopic maps may impose limits on how well visual information is encoded and remembered (Emrich et al., 2013; Sprague, Ester, & Serences, 2014).

The impact of presenting memoranda in different depth planes is less clear. Given that the retina encodes a 2-D projection of light coming from a complex 3-D environment, depth information must be indirectly inferred based on binocular cues like retinal disparity and monocular cues from pictorial depth indicators. In addition to the second-order nature of depth computations, there is also far less evidence of maplike 3-D spatial representations in visual cortex. However, a recent study suggests that there are topographic representations of depth encoded in some visual areas, so separation in 3-D may operate much like separation in 2-D (Finlayson, Zhang, & Golomb, 2017). In addition, studies of visual search suggest that 3-D structure may generally facilitate information processing. For example, visual-search performance is better when depth information is present, particularly when the 3-D structure of the display is kept constant across trials (McCarley & He, 2001). Visual-search performance is also substantially better when participants are searching for a combination of color and depth or motion and depth compared to searching for a combination of two visual features that are not separated in depth. This finding suggests that depth separation can facilitate the separate encoding of visual features (Nakayama & Silverman, 1986).

That said, the few previous studies that have directly investigated the effect of depth on VWM task performance have reported conflicting evidence, with some finding performance improvements and some finding performance decrements (Qian, Li, Wang, Liu, & Lei, 2017; Reeves & Lei, 2014; Xu & Nakayama, 2007). In addition, studies focusing on different aspects of information processing, such as selective attention, suggest that separating visual stimuli in depth might lead to impaired performance because encoding across different depth planes increases the total volume of 3-D space that participants must attentively monitor (Andersen, 1990; Andersen & Kramer, 1993; Atchley, Kramer, Andersen, & Theeuwes, 1997; Downing & Pinker, 1985; Enns & Rensink, 1990; Finlayson & Grove, 2015; Finlayson, Remington, Retell, & Grove,

2013; Theeuwes, Atchley, & Kramer, 1998). For instance, while attention tends to naturally spread across perceived 3-D surfaces, it is not as easy to divide attention between two 3-D surfaces (He & Nakayama, 1995). Similarly, separating memoranda in depth might hinder performance because of these limitations in attention. Thus, it remains unclear whether depth would be important in the same way as 2-D space for improving the separability of representations in working memory.

To test these alternative accounts, we examined the effects of 2-D in-plane and 3-D depth separation on memory precision (Experiment 1) and interactions between separation in depth and the number of remembered items (i.e., the set size of the memory array; Experiment 2). In Experiment 1, we found that separating items in depth improves memory performance in a manner similar to separating items in the 2-D plane. In Experiment 2, we found that the benefits of separating memoranda in depth were particularly evident in participants who were better able to perceive items in depth, and when participants had to remember a larger number of items. Together, these findings show that both 2-D in-plane and 3-D across-planes spatial separability improve VWM performance. Thus, performance benefits for items separated in the 2-D plane may extend to structured representations of the inferred 3-D layout of a visual scene, perhaps as a result of the recruitment of more retinotopically distinct neural resources.

## Experiment 1

### Methods

#### Participants

Thirty healthy volunteers (21 women, nine men; mean age [ $\pm$  standard error of the mean] = 20.87  $\pm$  0.53 years) from the University of California San Diego (UCSD) community participated in the experiment. All procedures were approved by the UCSD Institutional Research Board. All participants reported normal or corrected-to-normal vision without color-vision deficiency, and provided written informed consent. To ensure that all participants had stereovision, we prescreened for stereo blindness by asking all participants to look at a random-dot stereogram display through binocular goggles and then identify three different geometric shapes (a triangle, a square, and a circle). These shapes can be seen only if participants successfully fuse the images from the left and right eyes. All participants in this study correctly identified all three shapes. Participants were unaware of the purpose of the study and received course credit for their time.

Three participants were excluded from the analysis due to low performance (circular standard deviation of more than 45°).

#### Stimuli and procedure

Stimuli were rendered using virtual-reality goggles (Oculus DK2, Microsoft, Redmond, WA) with a resolution of 1,920  $\times$  1,080, at a 60-Hz refresh rate and a screen size of 12.6  $\times$  7.1 cm (subtending 90°  $\times$  60° visual angle). They were generated on a PC running Ubuntu (version 16.04) using MATLAB and the Psychophysics Toolbox (Brainard, 1997; Pelli, 1997). Participants were instructed to maintain fixation on a white central fixation dot (0.25° diameter) presented on a midgray background of 6.54 cd/m<sup>2</sup>. To aid ocular fusion and maintain stable and vivid depth perception, 16 gray circular placeholders (each 0.8° in diameter) were presented at evenly spaced intervals along an imaginary circle with a radius of 2.5°. The location of the placeholders in depth was either  $-0.1^\circ$  or  $0.1^\circ$ , based on retinal disparity. Depth was varied such that alternating pairs of placeholders had either a positive or a negative disparity (i.e., two close, then two far, then two close, etc.; see Figure 1). Memory-item colors were selected from a circle in CIE La\*b\* color space ( $L = 70$ ,  $a = 20$ ,  $b = 38$ , radius = 60). The two target colors were always 90°  $\pm$  10° apart along the circular color space. We opted to maintain this separation in color space so that the separability of the memory items in color space would remain relatively stable, allowing us to manipulate only 2-D and 3-D spatial separability across experimental conditions. The two memory targets were always presented either close in 2-D space (adjacent, with their centers 0.98° apart) or farther away (centers 2.78° apart), and they could be on the same or different depth planes. This produced four levels of 3-D (same vs. different) and 2-D (close vs. far) separation: same-close, different-close, same-far, and different-far. Note that the two memory targets were always presented in the same hemifield to maximize interitem competition (Alvarez & Cavanagh, 2005; Cohen et al., 2016; Störmer, Alvarez, & Cavanagh, 2014). No color calibration was done on the Oculus goggles. However, since the locations, sizes, and colors of memory items are consistent across all conditions, we believe that any error from calibration will affect all conditions equally. In general, the error introduced by the memory task itself is very large relative to any display properties; reliable data in such paradigms can even be obtained in continuous color-report tasks conducted in entirely uncontrolled settings (e.g., over the internet with all subjects using their own personal computer; Brady & Alvarez, 2015).

On each trial, two colored stimuli were presented for 150 ms and participants had to remember the color of

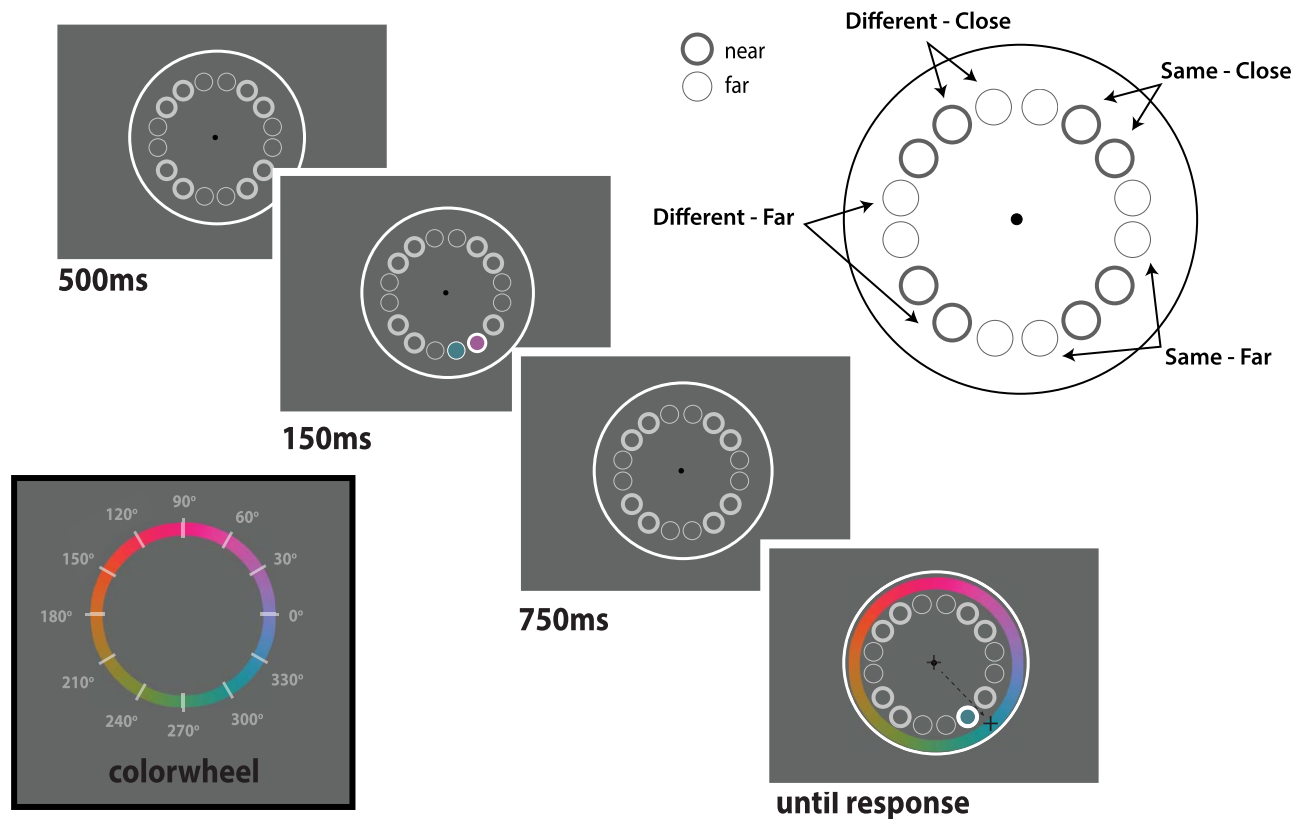


Figure 1. Each trial started with a 500-ms fixation period during which only the 16 placeholders were shown. Here, light and dark circles indicate placeholders on the far and near depth planes, respectively (this is only for visualization purposes—all placeholders were the same shade of gray in the actual experiment). Next, two memory targets were presented for 150 ms, followed by a 750-ms delay. After the delay, a color wheel was presented together with a cue outlining one of the previous target locations, and participants moved the cursor to report the hue previously shown at the cued location. The two target colors were presented in either the same or different depth planes in 3-D coordinates (same vs. different) and either close or far in 2-D space (see insert at top right). The lower left insert shows the color wheel that we used in the experiment.

both during a 750-ms delay period. After the delay, one of the two colors was probed by increasing the thickness of one of the placeholders. Together with the location probe, a color wheel ( $3^\circ$  radius from the center,  $0.5^\circ$  wide, randomly rotated on each trial) and a crosshair appeared over the fixation dot. Participants used the mouse to move the crosshair from the fixation dot to the hue on the color wheel that most closely resembled the color of the probed memory target (Wilken & Ma, 2004). The next trial started after participants clicked the mouse to record their response, and this procedure was repeated 96 times per experimental condition (384 trials in total, conditions randomly interleaved).

### Analyses

We generated a distribution of errors for each participant by computing the difference between the cued target color and the reported color ( $\text{reported}^\circ - \text{target}^\circ$ ) on each trial. To clearly visualize the shape of this error distribution and its relationship to the

nontarget color, we flipped the sign of the error such that the nontarget color was always  $90^\circ$  counterclockwise from the cued target (Figure 2). A commonly used mixture model (Bays et al., 2009; Zhang & Luck, 2008) was fitted to the error distribution under the assumption that responses reflect a mixture of responses to the target color, responses to the nontarget color, and random guesses. This model had four free parameters: the bias ( $b$ , in degrees) of the responses, the standard deviation ( $SD$ ) of the responses (both target and nontarget), the probability of swapping errors ( $s$ , in percent), and the guess rate ( $g$ , in percent; Bays, 2015; Bays et al., 2009; Zhang & Luck, 2008). It was fitted separately to data from each condition for each participant using Mem-Toolbox (Suchow, Brady, Fournie, & Alvarez, 2013). A repeated-measures analysis of variance was then performed to evaluate the impact of 2-D (near/far) and 3-D (same/different depth plane) spatial separation on the estimated model parameters.

It is important to note that the mixture model may have limitations (Schurgin et al., 2018); in particular, precision and guess rate may not be truly separable

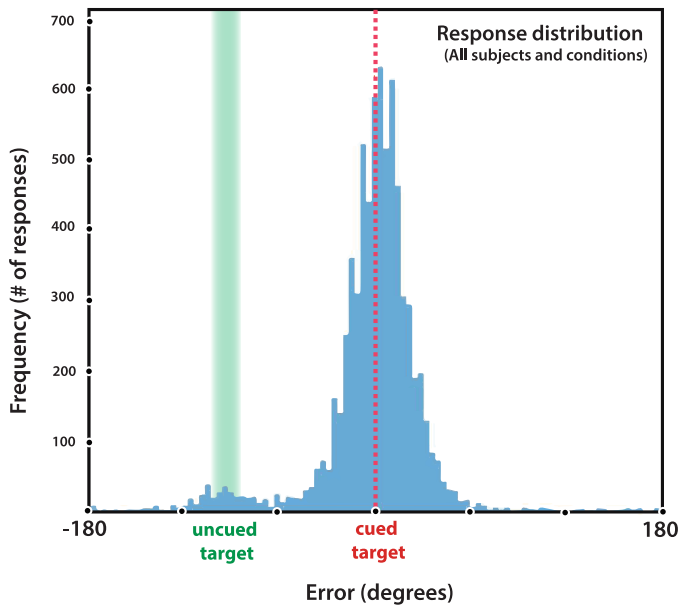


Figure 2. Results of Experiment 1 as a histogram of the responses centered around the target color, shown collapsed across all participants and conditions. The nontarget colors were aligned to approximately  $-90^\circ (\pm 10^\circ)$  relative to the target color by flipping the sign of responses on trials where the nontarget was  $+90^\circ (\pm 10^\circ)$  relative to the target (note that the width of the shaded green area reflects the  $\pm 10^\circ$  jitter in the uncued target color). Swap errors are apparent from the small bump centered on the nontarget color.

parameters. However, we opted to use the mixture model in this particular experiment because it allowed us to account for systematic biases and for responses to nontargets (swap errors), which are difficult to account for without using a model of the response distribution. For example, without explicit accounting for swap errors, nontarget responses would be treated as  $90^\circ$  errors even though they were actually accurate responses to the nontarget color. However, to check that our results were not dependent on the details of the mixture model, we also performed a post hoc analysis where we developed a nonparametric procedure to quantify memory precision while taking systematic biases and swap errors into account: First, we computed the error (in degrees) of all responses that were centered around the target and nontarget colors (i.e., including responses to nontarget colors as precise responses). Then, in an effort to attenuate the effect of systematic biases, we computed the mean absolute error within  $\pm 60^\circ$  from the peak (mode) of each error-response distribution (i.e., target and nontarget distributions). This allowed us to nonparametrically examine errors without any strong assumptions about the separability of the guess rate and precision parameters of a mixture model.

## Results

Responses were more precise (lower mixture-model *SD*) both when the two memoranda were separated by a greater distance in 2-D spatial position (near/far),  $F(1, 26) = 4.921$ ,  $p = 0.036$ , and when the two memoranda were presented on different depth planes (same/different planes),  $F(1, 26) = 5.677$ ,  $p = 0.025$ , with no interaction between these factors,  $F(1, 26) = 0.06$ ,  $p = 0.808$  (Figure 3A). As shown in Figure 3B, there was a consistent bias such that responses were repelled slightly but consistently away from the nontarget color,  $t(1, 26) = 5.81, 6.63, 6.47$ , and  $7.77$ , respectively, for same-close, different-close, same-far, and different-far, all  $ps < 0.0001$ . However, there was no difference in the magnitude of this bias as a function of separation in 2-D,  $F(1, 26) = 0.002$ ,  $p = 0.965$ , or in 3-D,  $F(1, 26) = 1.377$ ,  $p = 0.251$ , and no interaction between these factors,  $F(1, 26) = 0.983$ ,  $p = 0.331$ . The probability of swapping (i.e., nontarget reports; Figure 3C) did not depend on whether the items were spatially close or far away from each other in 2-D space,  $F(1, 26) = 1.633$ ,  $p = 0.213$ , and there was a nonsignificant trend toward more swap errors when targets were presented on different depth planes,  $F(1, 26) = 3.211$ ,  $p = 0.085$ . No interaction was observed,  $F(1, 26) = 1.889$ ,  $p = 0.181$ . There were also no differences in guess rates estimated by the mixture model across conditions—separation in 2-D:  $F(1, 26) = 0.008$ ,  $p = 0.93$ ; separation in 3-D:  $F(1, 26) = 1.481$ ,  $p = 0.235$ ; interaction:  $F(1, 26) = 0.366$ ,  $p = 0.55$  (Figure 3D).

The quantitative results from this mixture modeling match with the qualitatively observable shapes of the kernel density plots for each condition (Figure 3A–3D vs. 3E, computed using a Gaussian kernel with a standard deviation of  $4^\circ$ ), and the nonparametric analysis of response precision yielded comparable results: The average absolute error around the target was higher when two items were separated both in 2-D,  $F(1, 26) = 6.66$ ,  $p = 0.016$ , and in 3-D,  $F(1, 26) = 6.40$ ,  $p = 0.018$ , and there was no interaction,  $F(1, 26) = 0.46$ ,  $p = 0.505$ .

To evaluate statistical power in our study, we performed a post hoc bootstrapping analysis in which we systematically varied the number of participants. We resampled with replacement data from different numbers of participants, ranging from two to 27, and on each resample we computed the mean differences between conditions. This process was then repeated 1,000 times. On each iteration, we did the same analysis of both the parameters from the mixture model and the nonparametric mean absolute error, and found that both analyses reached stable statistical significance (two-sided  $p < 0.05$ ) with a minimum of 20 participants.

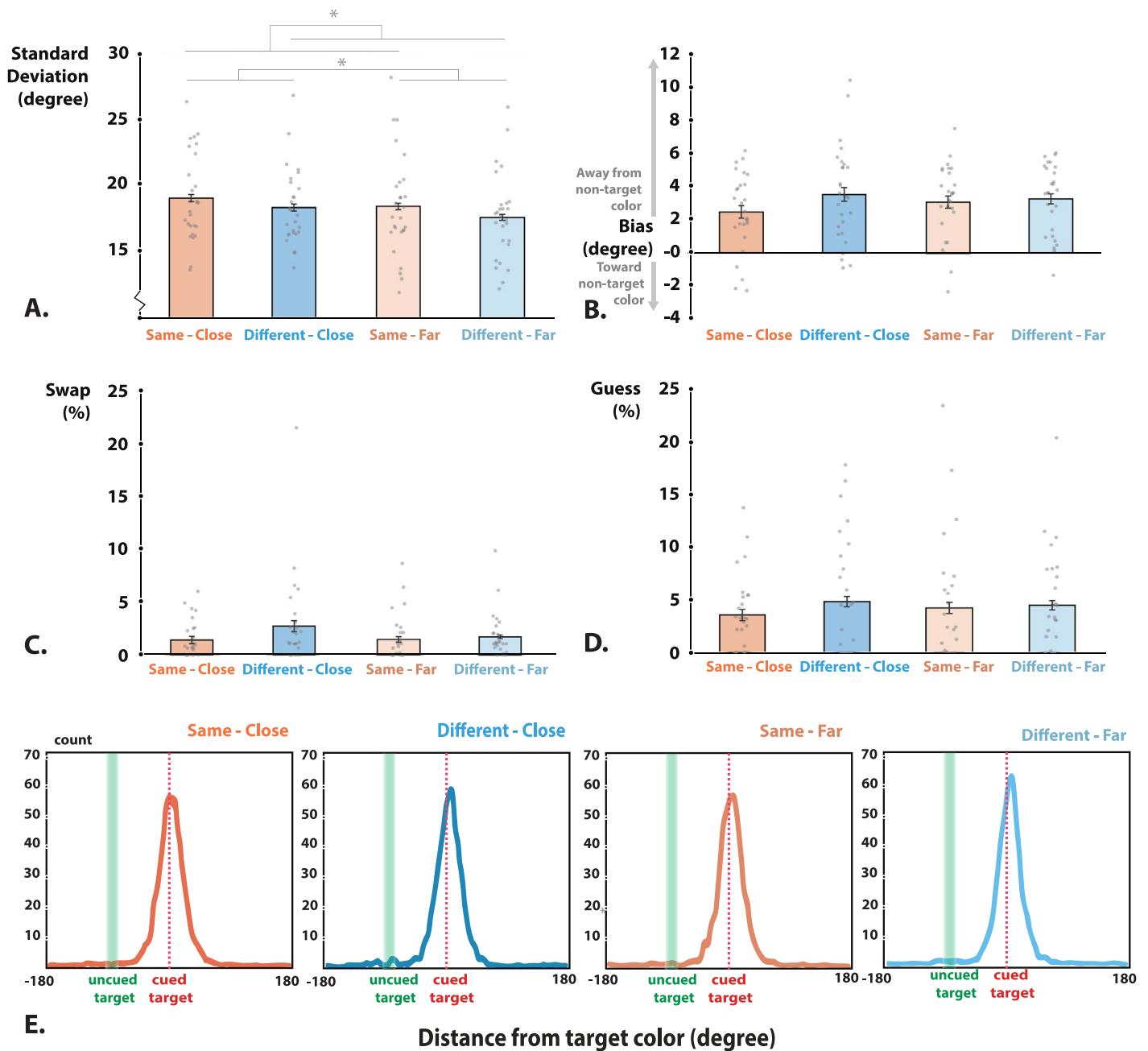


Figure 3. Results of Experiment 1 in terms of the parameters from mixture modeling. (A) The standard deviations are lower when two memory items are spatially far away or when they are on different depth planes (lower standard deviation is associated with higher precision).  $*p < 0.05$ . (B) There are systematic biases away from the nontarget color in all conditions but no significant differences in biases between conditions. (C) There are no significant differences in swap error rate, nor in (D) guess rate. (E) Four kernel density plots of group-level error responses of each condition centered around the target color (from left: same-close, different-close, same-far, and different-far). The shapes of the distributions qualitatively agree with the parameters from the model. Error bars (in A, B, and C,) represent  $\pm 1$  standard error of the mean.

Together these results suggest that spatial separability both within and between different depth planes is associated with higher precision memories in VWM. Importantly, no effects of spatial separability were found on any of the other parameters, suggesting that it is the memory strength that improves once items are separated in either 2-D or 3-D space.

Finally, note that the bias we observed in the target responses was always positive, or away from the nontarget, which is consistent with previous studies showing repulsion biases away from other task-relevant items (Bae & Luck, 2017; Golomb, 2015; Marshak & Sekuler, 1979; Rademaker, Bloem, De Weerd, & Sack, 2015; Rauber & Treue, 1998; Scocchia, Cicchini, &

Triesch, 2013). Interestingly, one study that examined repulsion bias as a function of color similarity between items (Golomb, 2015) showed repulsion biases only when items were close in feature space—specifically less than 60° apart in feature space—while attraction biases were reported when memoranda were more than 60° apart in feature space. However, in the current study we observed repulsion biases even with colors separated by 90° in feature space. Numerous aspects of the current task differed from that previous work (e.g., number of memory items, encoding time, delay time), and many of these factors could affect whether repulsion or attraction is observed in the data and account for the differences between findings.

## Experiment 2

The results from Experiment 1 suggest that separating memoranda within and between depth planes increases memory precision, presumably because interference between the items is reduced. Here we examine the effects of depth on VWM capacity, focusing on the ways it might improve attentional filtering. Studies have shown that the number of items that people can hold in memory with high fidelity may decrease once the number of items to be remembered is large and difficult to manage. For example, one person might be capable of remembering four items with a high degree of fidelity when there are only four to be remembered. However, that same person might remember fewer than four items with a high degree of fidelity when there are 12 memoranda to retain (Cowan & Morey, 2006; Cowan, Morey, AuBuchon, Zwilling, & Gilchrist, 2010; Cusack, Lehmann, Veldsman, & Mitchell, 2009; Linke, Vicente-Grabovetsky, Mitchell, & Cusack, 2011; Vogel, McCollough, & Machizawa, 2005). This phenomenon has usually been attributed to a failure of attentional filtering, as trying to store everything in the display may have negative consequences. Previous work has shown that spatial location can aid attentional filtering (Vogel et al., 2005). Therefore, we hypothesized that separating items in depth might also aid attentional filtering. In particular, we predicted that once participants have a large number of items to remember and therefore must rely on attentional filtering to select a subset of items to represent with high fidelity, separation in depth should promote a higher memory capacity. Alternatively, it is possible that increasing the number of memory items in a 3-D display might lead to poorer overall performance due to an increased demand to distribute spatial attention across a larger volume of space. To test these accounts, we manipulated memory set size across a range from two to 12 items. We also independently

assessed each participant's ability to exploit stereo depth cues so that we could evaluate the relationship between the salience of depth information and its impact on VWM capacity across participants.

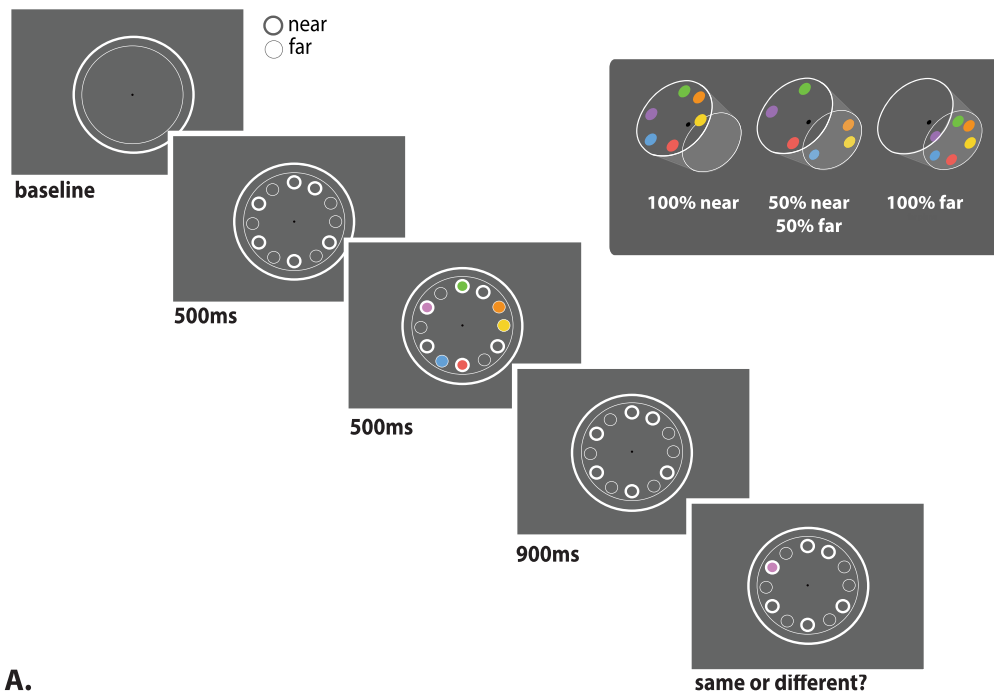
## Methods

### Participants

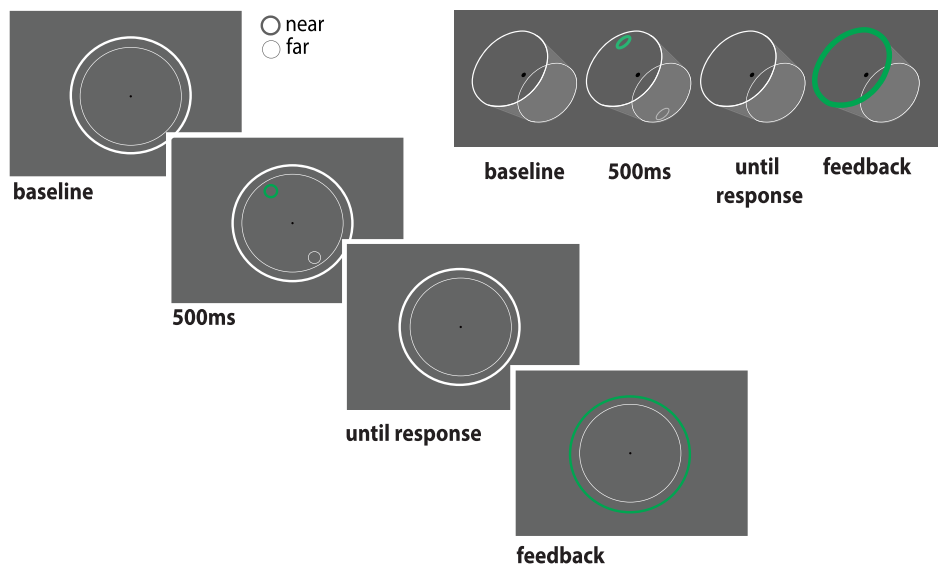
A new set of 22 healthy volunteers (14 women, eight men; mean age =  $19.67 \pm 0.45$  years) from the UCSD community participated in the experiment. All procedures were approved by the UCSD Institutional Research Board. All participants reported normal or corrected-to-normal vision without color-vision deficiency and provided written informed consent. Participants were unaware of the purpose of the study and received course credits or monetary compensation for their time (\$10/hr). All participants passed the same stereovision test used in Experiment 1, and none were excluded.

### Stimuli and procedure

Unless otherwise mentioned, stimulus generation and presentation were identical to Experiment 1. The main VWM task in Experiment 2 (Figure 4A) used a delayed-match-to-sample paradigm. At the beginning of each trial, 12 placeholders were presented (each 1° in diameter, presented at 2.5° from fixation) for 500 ms. The depth separation of the placeholders was experimentally manipulated: On 50% of trials, they were all presented on the same depth plane (all on the near plane on 25%, all on the far plane on another 25%)—the same-depth condition. On the remaining 50% of trials, half of the placeholders were on the near plane and the other half were on the far plane—the different-depths condition. Next, two, four, six, eight, or 12 colored memory targets were briefly presented (500 ms) at a random subset of the 12 placeholders, with the restriction that in the different-depths condition half of the items were assigned to near and the other half to far placeholders (12 stimuli were shown in every placeholder). Colors were randomly chosen from a set of 12 unique colors. After a 900-ms delay, a single test color was presented at one of the memory-target locations, either matching or not matching the target color previously shown at that location. Participants indicated match or nonmatch by pressing the X or C key, respectively, with matches occurring on 50% of trials and nonmatches created by placing one of the other remembered items from the initial display in the test location. For each participant, we collected 80 trials for each set-size (2, 4, 6, 8, and 12) and depth condition (same vs. different depths), leading to 800 total trials. Participants performed 10 blocks of 80 trials each, with



**A.**



**B.**

Figure 4. Experimental procedure for Experiment 2. (A) In this single-probe change-detection paradigm, each trial started with the presentation of 12 placeholders. Placeholders could have one of three possible depth relationships: all on the near depth plane, all on the far depth plane, or half on the near and the other half on the far depth plane. After 500 ms, two, four, six, eight, or 12 colored memory items were presented for 500 ms, followed by a 900-ms delay period. Next, a single test item was presented at a location previously occupied by one of the memory items, and participants indicated whether the color of the test was the same as or different from the color of the memory target previously shown at that location. (B) The independent depth-discrimination task. On each trial, two placeholders briefly appeared, each on a different depth plane. Participants indicated whether the target (in green) was on the near or far plane. Performance on this task was used as an indicator of how well participants could perceive depth using our stereo-display setup.



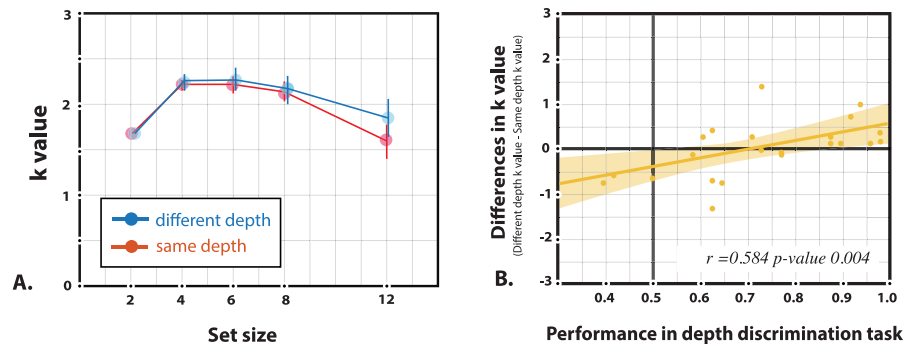


Figure 5. Main results of Experiment 2. (A) Visual-working-memory capacity (Cowan's  $k$ ) as a function of set size. There were no differences in capacity when memory items were displayed on planes at the same (red) or different (blue) depths. Observed changes in  $k$  as a function of set size are consistent with previous studies (Cowan & Morey, 2006). (B) The impact of depth separation (on the y-axis) was calculated by taking the capacity  $k$  for items presented on different depth planes minus the  $k$  for items presented on the same depth plane. Thus, larger numbers indicate a larger benefit of presenting items separated in depth. The ability of participants to discriminate the two depth planes in our experimental setup (on the x-axis) was positively correlated with the benefits they gained from items presented on different depth planes. Shaded regions indicate  $\pm 1$  standard error of the mean.

each block lasting  $\sim 5$  min. Note that using a delayed-match-to-sample paradigm required less time per trial than continuous report and thus allowed us to quickly evaluate memory performance across five set sizes for items on the same and different depth planes.

To enable us to evaluate how well participants could perceive memoranda presented on the two different depth planes, participants also completed a 48-trial depth-discrimination task (Figure 4B) prior to participating in the main task. During this independent task, two placeholders were presented for 500 ms, with one on the near plane and the other on the far plane (with respect to fixation). The location of the two placeholders was chosen at random from the 12 possible locations used in the main task. Participants had to indicate whether a target (specified by a green circle outline) was on the near or far plane. The ability of each participant to accurately identify the correct depth plane in this task was used to predict the benefits of the depth information during the VWM task.

### Analyses

We estimated each participant's VWM capacity using a standard measure appropriate for single-probe change detection—Cowan's  $k$  (Cowan, 2010; Pashler, 1988)—as follows:

$$k = (\text{hit rate} - \text{false alarm}) \times \text{set size}.$$

As in Experiment 1, repeated-measures analyses of variance were used for the main analyses. Additionally, the impact of a participant's ability to perceive the stimuli in depth (measured with the independent depth-discrimination task) on performance during the VWM task was assessed using correlational analyses.

### Results

There was a significant main effect of set size on observed  $k$  values,  $F(4, 84) = 5.26$ ,  $p < 0.001$  (Figure 5A), such that estimates of capacity were lower for very small and very large set sizes; a linear fit failed to capture a significant amount of variance,  $F(1, 215) = 0.59$ ,  $p = 0.44$ , while adding a quadratic significantly improved the fit,  $F(3, 215) = 3.81$ ,  $p = 0.011$ . However, there was no effect of depth condition,  $F(1, 21) = 0.018$ ,  $p = 0.895$ , and no Set size  $\times$  Depth condition interaction,  $F(4, 84) = 0.107$ ,  $p = 0.98$ . While this may suggest that presenting memory items on the same versus different depth planes did not affect memory capacity, we found a positive correlation between depth-discrimination ability (as indexed during the independent depth-discrimination task) and the impact of separation in depth (as manipulated in the main VWM task). Specifically, participants with better stereo depth perception showed a larger performance benefit when items were presented on different depth planes (Pearson's  $r = 0.58$ ,  $p = 0.004$ ; Figure 5B), and this correlation was still significant when participants with negative  $k$  values were excluded from the analysis (Pearson's  $r = 0.55$ ,  $p = 0.012$ ). This effect was systematically related to set size, such that correlations grew stronger as set size increased (Figure 6, bottom row)—set size 2:  $r < 0.0001$ ,  $p = 0.99$ ; set size 4:  $r = -0.05$ ,  $p = 0.81$ ; set size 6:  $r = 0.38$ ,  $p = 0.08$ ; set size 8:  $r = 0.42$ ,  $p = 0.05$ ; set size 12:  $r = 0.54$ ,  $p = 0.008$ .

Importantly, the correlations between the depth-discrimination task and VWM performance were found selectively in the 3-D condition (Pearson's  $r = 0.49$ ,  $p = 0.05$ ) but were not found in the 2-D condition (Pearson's  $r = 0.05$ ,  $p = 0.80$ ). The correlation analyses after excluding two subjects with negative average  $k$

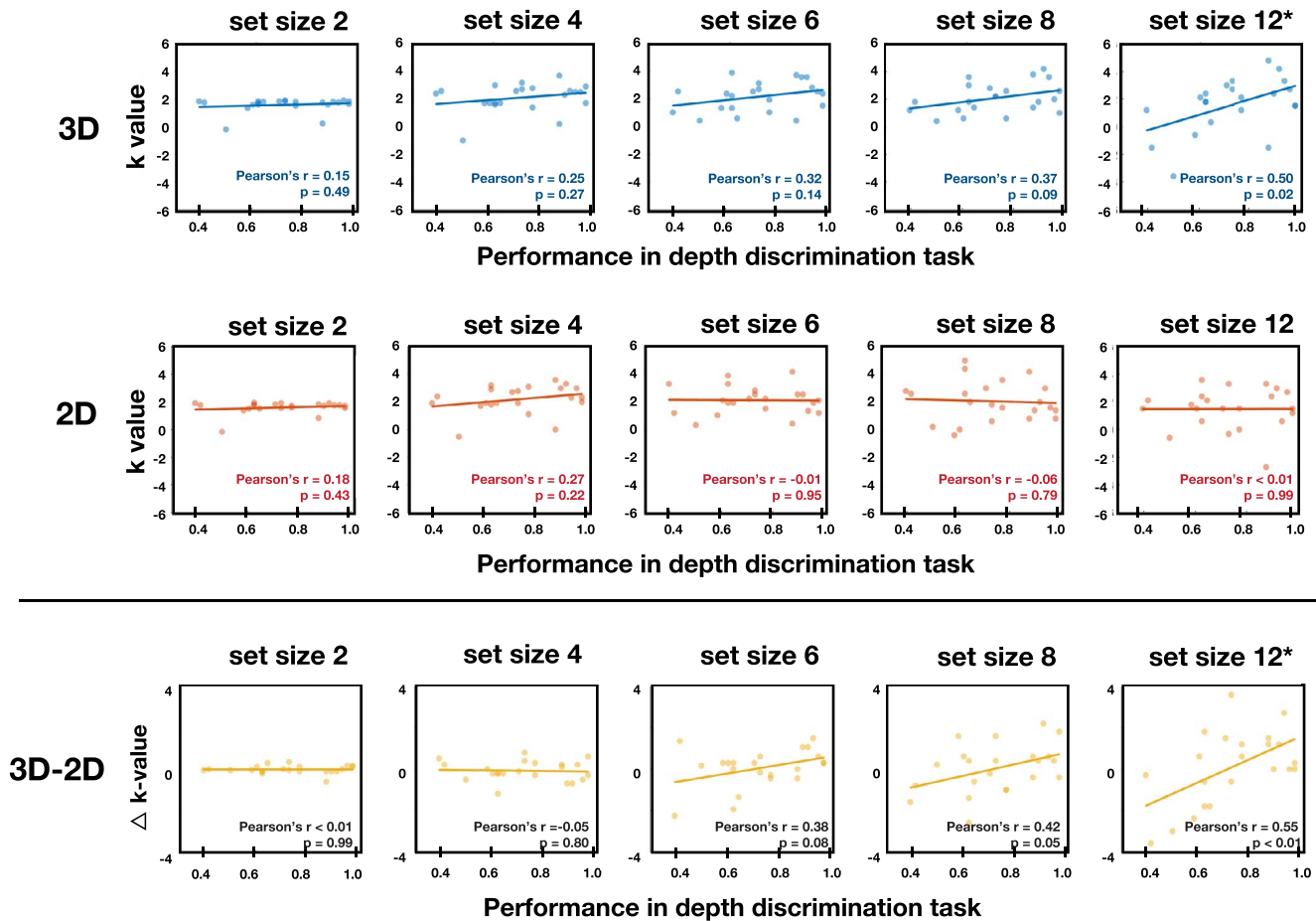


Figure 6. The degree of positive correlation between depth-discrimination ability (on the x-axis) and performance on the visual-working-memory task (on the y-axis). Participants who performed better on the depth-discrimination task also performed better on the visual-working-memory task at larger set sizes, but only when the memoranda were on different depth planes (upper row). There was no correlation between performance on the depth-discrimination task and on the visual-working-memory task when the memoranda were in the same depth plane (middle row). The benefit associated with having the memoranda separated into different depth planes (difference in  $k$  value on the y-axis) grew stronger as set size increased (bottom row in panels).

values found similar results (3-D: Pearson's  $r = 0.49$ ,  $p = 0.028$ ; 2-D: Pearson's  $r = -0.008$ ,  $p = 0.97$ ). We ran a dependent correlation test and found a significant difference between the 2-D and 3-D correlations,  $t = 3.08$ ,  $p = 0.01$ , showing that the 3-D correlations were reliably higher than in the 2-D condition. This indicates that the correlation was not related to differences in general arousal or motivation (Figure 6). We believe that the effect is robust given that these correlations grow monotonically stronger as set size increases. To ensure that this analysis had enough power, we did a bootstrapping analysis in which we resampled data from a different number of participants (between five and 22) with replacement 1,000 times (just as we did in Experiment 1). We found stable positive correlations (more than 97.5% of the simulations had positive correlations; equal to two-sided  $p < 0.05$ ) when there were at least 10 participants included.

As an alternate means of assessing the data, we sorted participants into two groups based on a median split of their depth-discrimination ability as assessed using the independent task (Figure 7). We found a main effect of set size,  $F(4, 80) = 5.22$ ,  $p < 0.001$ , but not of depth plane,  $F(1, 20) = 0.03$ ,  $p = 0.87$ . There was also a significant two-way interaction such that separation in depth led to improved performance only for those subjects who performed well on the independent depth-discrimination task,  $F(1, 20) = 10.95$ ,  $p = 0.004$ . Performance on the depth-discrimination task was not associated with an overall change in VWM performance levels collapsed across set size and condition,  $F(1, 20) = 0.79$ ,  $p = 0.39$ , suggesting that the two groups of subjects were equally motivated to perform the task. Nevertheless, there was a three-way interaction such that participants who performed well on the independent depth task showed the benefit of depth at larger set size,  $F(4, 80) = 3.622$ ,  $p = 0.009$ .

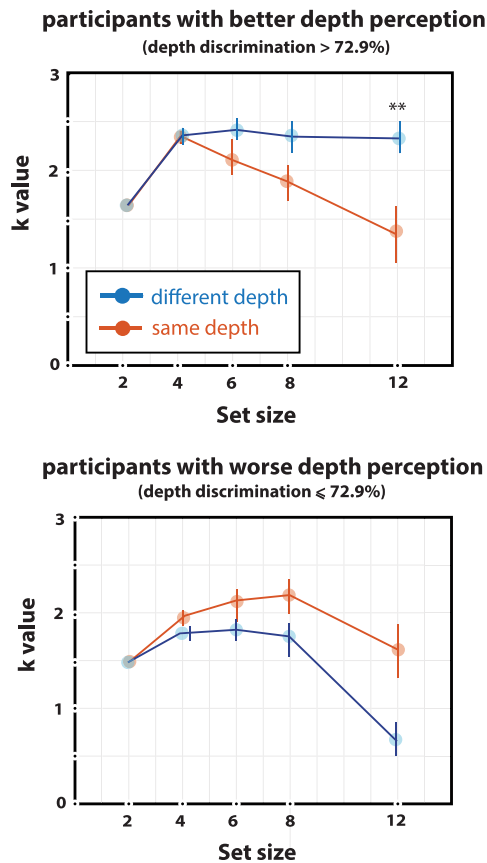


Figure 7. Participants who exhibited better depth discrimination (upper panel), based on a median split of performance in the independent depth-discrimination task, benefited more from the presence of depth information, particularly at high set sizes.  $**p < 0.01$ . The error bars represent  $\pm 1$  standard error of the mean. For participants who exhibited worse depth discrimination (lower graph), the  $k$  value appeared to be lower when memoranda were on different depth planes, but this did not reach significance. Note that the performance from both groups was comparable when the memoranda were on the same depth plane (compare red lines between the two panels).

To follow up on these findings, we also performed post hoc tests separately on data within the low and high depth discriminators. We found that the high depth discriminators did better on the VWM task when the items were separated in depth—main effect:  $F(1, 11) = 6.79$ ,  $p = 0.024$ —especially with larger set sizes—interaction:  $F(4, 44) = 3.53$ ,  $p = 0.014$ . This indicates that participants with better depth perception ( $>72.9\%$  accuracy) performed better on different-depths displays, but only at larger set sizes (Figure 7, top panel)—set size 2:  $t(1, 11) = -0.25$ ,  $p = 0.81$ ; set size 4:  $t = 0.06$ ,  $p = 0.96$ ; set size 6:  $t = 1.83$ ,  $p = 0.09$ ; set size 8:  $t = 1.44$ ,  $p = 0.18$ ; set size 12:  $t = 2.78$ ,  $p = 0.02$ . For the low depth discriminators there was a small opposite trend such that performance was lower when memoranda were in different depth planes. However, the analysis of

variance did not reveal a significant main effect of separation in depth,  $F(1, 9) = 4.439$ ,  $p = 0.064$ , nor an interaction,  $F(4, 36) = 1.052$ ,  $p = 0.394$ . And post hoc paired  $t$  tests were also nonsignificant (Figure 7, bottom panel)—set size 2:  $t(1, 9) = -0.35$ ,  $p = 0.73$ ; set size 4:  $t = -1.35$ ,  $p = 0.21$ ; set size 6:  $t = -0.78$ ,  $p = 0.46$ ; set size 8:  $t = -1.14$ ,  $p = 0.29$ ; set size 12:  $t = -1.63$ ,  $p = 0.14$ .

We also performed post hoc tests separately on data from same- and different-depth conditions. Importantly, there was an interaction between low and high depth discriminators and set size when the memoranda were on different planes,  $F(4, 80) = 2.87$ ,  $p = 0.028$ , but not when they were on the same plane,  $F(4, 80) = 0.75$ ,  $p = 0.564$ , indicating that the benefits of better depth perception were restricted to trials where the memory load was high and memoranda were presented in separate depth planes. Moreover, the lack of an effect of depth-perception ability on performance in the same-depth condition further suggests that differences in overall motivation between the two groups of participants cannot account for the observed differences in the different-depths condition.

## Discussion

Perceiving the world in 3-D is a seemingly effortless endeavor, and depth information is fundamental to perceptual organization of the visual world into objects and surfaces, as well as guiding motor interactions with objects in the environment. However, the manner in which the visual system represents in-plane 2-D information versus 3-D depth information is fundamentally different. First, depth information must be indirectly inferred based on operations applied to the 2-D input provided by the projection of light onto the retina. Thus, depth is a second-order feature of visual representation that is indirectly constructed from a set of binocular and monocular cues. Second, the visual system is organized such that ordinal information about the 2-D layout of a visual scene is preserved: Stimuli that are closer to each other in the world are represented by neurons that are closer to each other in the retina and in later visual areas. In contrast, the extent of topographic representations of depth in visual cortex is not well understood, with only a few recent studies suggesting that a structured layout of depth exists in some visual areas (Finlayson et al., 2017). Here we show that separating memoranda in both the 2-D plane and 3-D depth improves VWM performance, consistent with the idea that separating stimuli in depth attenuates interitem competition and interference which affects how people perceive the display (Andersen, 1990; Finlayson & Golomb, 2016; Kooi, Toet, Tripathy, & Levi, 1994; Lehmkuhle & Fox, 1980;

Papathomas, Feher, Julesz, & Zeevi, 1996). This is also in line with evidence that people remember real-world 3-D objects better than drawings or photographs of the same objects, even when retinal images are roughly matched (Snow, Skiba, Coleman, & Berryhill, 2014). Furthermore, separating memoranda in depth had the biggest impact on performance when set size increased, suggesting that at least some participants were able to exploit this additional 3-D spatial information to help encode and maintain distinct representations of remembered items.

Previous work has produced mixed results regarding the impact of depth on VWM. For example, two recent studies using a change-detection task did not find any effect of separating memoranda in depth using a display in which all items were presented simultaneously (Qian et al., 2017; Reeves & Lei, 2014). An earlier study also found no benefits of depth using a simultaneous display, but did find that participants had a higher VWM capacity under stereoscopic viewing conditions when each item was presented sequentially on a different depth plane (Xu & Nakayama, 2007). The authors of this latter study hypothesized that perceiving items separated in depth might be inherently more difficult in a simultaneous display, as participants need to attend more than one depth plane at a time—in sequential displays this is presumably no longer an issue, unveiling the benefits of separation in depth. Interestingly, that same study showed that separation in depth had a benefit above and beyond other grouping cues, like changing the configuration of the memoranda by grouping subsets of memoranda into squares or circles. However, in everyday life we perceive depth information in stable and whole scenes, not in sequence. Because sequential presentation of depth information is one step removed from real-world conditions, it thus remains unclear from this previous work whether separation in depth yields any benefit without separation in time.

One alternative explanation for previous results which did not find a benefit to depth when using simultaneous displays is that participants simply differ in terms of how well they perceive the depth cues used in the experimental displays. In our Experiment 2, we independently measured individual differences in depth perception and found a clear benefit for separating memoranda in depth within the group of participants who were better able to exploit stereo cues to support depth perception. It is important to note that our depth-discrimination task required participants to be able to rapidly acquire depth information in order to accurately parse the array. Thus, even though all of the participants passed a basic stereovision screening test, there were still large individual differences in how efficiently they perceived depth information at the relatively brief exposure duration (i.e., 500 ms) used in

the depth-perception and VWM tasks. For example, participants who have stereovision but did poorly on the depth-perception task might not be able to rapidly switch their attention between depth planes (or not be able to simultaneously attend to both depth planes), resulting in relatively worse performance in the 3-D condition of the VWM task. The results from Experiment 2 also showed greater benefits of separation in depth at larger set sizes, consistent with the idea that separation in depth attenuates interitem competition and possibly improves attentional filtering. As visual attention (the ability to selectively process visual information) and VWM (the ability to retain visual information) are related cognitive mechanisms, one possibility is that the separation of items in depth affects how visual attention is distributed (e.g., sequential focal attention rather than simultaneous more distributed attention). Consequently, interference (and thus error) could be reduced, the difference between items amplified (two colors seen or remembered as more different; e.g., Finlayson & Golomb, 2016), and the relative position of items partially lost (more swap errors, e.g., mean nontarget responses of 19% vs. 4% in sequential vs. simultaneous display; Gorgoraptis, Catalao, Bays, & Husain, 2011).

It remains an open question to what extent our results arise from differences in binocular disparity per se, differences in perceived depth, or more general properties of surface perception (e.g., Nakayama, He, & Shimojo, 1995) regardless of the cues that give rise to such surfaces. Some work has suggested that perceptual benefits in related tasks are a result of binocular disparity rather than depth (Finlayson & Golomb, 2016), whereas many recognition tasks seem to largely benefit from coherent surface organization rather than binocular disparity (Nakayama, Shimojo, & Silverman, 1989). Future research will be needed to dissociate these different factors and their respective influences on VWM performance.

In summary, the present results demonstrate that separating memoranda in depth improves visual working memory. In Experiment 1, we show that separation in depth benefits VWM on a scale similar to separating memoranda in 2-D. The similarity of these depth effects to effects observed with 2-D space is particularly interesting given that spatial and depth information are fundamentally different, with 2-D information encoded directly at the retina while 3-D information needs to be indirectly inferred based on binocular and monocular cues. In Experiment 2, we show further that separation in depth confers the largest benefits when participants are better at exploiting stereo depth cues and when interitem competition is highest due to larger set sizes. Together, these observations suggest that interitem interference can occur after the computation of second-order properties

of the visual scene and not just at the level of retinotopically organized representations reflecting 2-D in-plane separation. Showing items at varying depths may thus confer an important benefit to behavioral performance in psychophysical tasks.

*Keywords:* visual working memory, memory biases, depth perception

## Acknowledgments

This work was supported by NEI R01-EY025872 and a James S. McDonnell Foundation Scholar Award to JTS, a Thai Red Cross Society grant to CC, by the European Union's Horizon 2020 research and innovation program under Marie Skłodowska-Curie Grant Agreement No. 743941 to RLR, and by NSF CAREER Award No. BCS-1653457 to TFB.

Commercial relationships: none.

Corresponding author: Chaipat Chunharas.

Email: chaipat.c@chula.ac.th.

Address: Department of Medicine, King Chulalongkorn Memorial hospital, Bangkok, Thailand.

## References

- Alvarez, G. A., & Cavanagh, P. (2005). Independent resources for attentional tracking in the left and right visual hemifields. *Psychological Science, 16*(8), 637–643.
- Andersen, G. J. (1990). Focused attention in three-dimensional space. *Perception & Psychophysics, 47*(2), 112–120.
- Andersen, G. J., & Kramer, A. F. (1993). Limits of focused attention in three-dimensional space. *Perception & Psychophysics, 53*(6), 658–667.
- Atchley, P., Kramer, A. F., Andersen, G. J., & Theeuwes, J. (1997). Spatial cuing in a stereoscopic display: Evidence for a “depth-aware” attentional focus. *Psychonomic Bulletin & Review, 4*(4), 524–529.
- Baddeley, A. (1986). Working memory, reading and dyslexia. In E. Hjelmquist & L.-G. Nilsson (Eds.), *Advances in Psychology* (Vol. 34, pp. 141–152). Amsterdam, the Netherlands: North-Holland.
- Bae, G.-Y., & Luck, S. J. (2017). Interactions between visual working memory representations. *Attention, Perception & Psychophysics, 79*(8), 2376–2395, <https://doi.org/10.3758/s13414-017-1404-8>.
- Bays, P. (2015). Evaluating and excluding swap errors in analogue report. *Journal of Vision, 15*(12): 675, <https://doi.org/10.1167/15.12.675>. [Abstract]
- Bays, P. M. (2016). Evaluating and excluding swap errors in analogue tests of working memory. *Scientific Reports, 6*, 19203.
- Bays, P. M., Catalao, R. F. G., & Husain, M. (2009). The precision of visual working memory is set by allocation of a shared resource. *Journal of Vision, 9*(10):7, 1–11, <https://doi.org/10.1167/9.10.7>. [PubMed] [Article]
- Bays, P. M., Gorgoraptis, N., Wee, N., Marshall, L., & Husain, M. (2011). Temporal dynamics of encoding, storage, and reallocation of visual working memory. *Journal of Vision, 11*(10):6, 1–15, <https://doi.org/10.1167/11.10.6>. [PubMed] [Article]
- Bays, P. M., & Husain, M. (2008, August 8). Dynamic shifts of limited working memory resources in human vision. *Science, 321*(5890), 851–854.
- Bays, P. M., Wu, E. Y., & Husain, M. (2011). Storage and binding of object features in visual working memory. *Neuropsychologia, 49*(6), 1622–1631.
- Brady, T. F., & Alvarez, G. A. (2011). Hierarchical encoding in visual working memory: Ensemble statistics bias memory for individual items. *Psychological Science, 22*(3), 384–392.
- Brady, T. F., & Alvarez, G. A. (2015). Contextual effects in visual working memory reveal hierarchically structured memory representations. *Journal of Vision, 15*(15):6, 1–24, <https://doi.org/10.1167/15.15.6>. [PubMed] [Article]
- Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision, 10*(4), 433–436.
- Cohen, M. A., Rhee, J. Y., & Alvarez, G. A. (2016). Limits on perceptual encoding can be predicted from known receptive field properties of human visual cortex. *Journal of Experimental Psychology: Human Perception and Performance, 42*(1), 67–77.
- Conway, A. R. A., Cowan, N., Bunting, M. F., Theriault, D. J., & Minkoff, S. R. B. (2002). A latent variable analysis of working memory capacity, short-term memory capacity, processing speed, and general fluid intelligence. *Intelligence, 30*(2), 163–183.
- Cowan, N. (2010). The magical mystery four: How is working memory capacity limited, and why? *Current Directions in Psychological Science, 19*(1), 51–57.
- Cowan, N., & Morey, C. C. (2006). Visual working memory depends on attentional filtering. *Trends in Cognitive Sciences, 10*(4), 139–141.
- Cowan, N., Morey, C. C., AuBuchon, A. M., Zwilling,

- C. E., & Gilchrist, A. L. (2010). Seven-year-olds allocate attention like adults unless working memory is overloaded. *Developmental Science*, *13*(1), 120–133.
- Cusack, R., Lehmann, M., Veldsman, M., & Mitchell, D. J. (2009). Encoding strategy and not visual working memory capacity correlates with intelligence. *Psychonomic Bulletin & Review*, *16*(4), 641–647.
- D'Esposito, M., & Postle, B. R. (2015). The cognitive neuroscience of working memory. *Annual Review of Psychology*, *66*, 115–142.
- Downing, C., & Pinker, S. (1985). *Attention and performance*. Mahwah, NJ: Earlbaum.
- Emrich, S. M., & Ferber, S. (2012). Competition increases binding errors in visual working memory. *Journal of Vision*, *12*(4):12, 1–16, <https://doi.org/10.1167/12.4.12>. [PubMed] [Article]
- Emrich, S. M., Riggall, A. C., Larocque, J. J., & Postle, B. R. (2013). Distributed patterns of activity in sensory cortex reflect the precision of multiple items maintained in visual short-term memory. *The Journal of Neuroscience*, *33*(15), 6516–6523.
- Engel, S. A., Glover, G. H., & Wandell, B. A. (1997). Retinotopic organization in human visual cortex and the spatial precision of functional MRI. *Cerebral Cortex*, *7*(2), 181–192.
- Engle, R. W., Tuholski, S. W., Laughlin, J. E., & Conway, A. R. (1999). Working memory, short-term memory, and general fluid intelligence: A latent-variable approach. *Journal of Experimental Psychology: General*, *128*(3), 309–331.
- Enns, J. T., & Rensink, R. A. (1990). Sensitivity to three-dimensional orientation in visual search. *Psychological Science*, *1*(5), 323–326.
- Finlayson, N. J., & Golomb, J. D. (2016). Feature-location binding in 3D: Feature judgments are biased by 2D location but not position-in-depth. *Vision Research*, *127*, 49–56.
- Finlayson, N. J., & Grove, P. M. (2015). Visual search is influenced by 3D spatial layout. *Attention, Perception & Psychophysics*, *77*(7), 2322–2330.
- Finlayson, N. J., Remington, R. W., Retell, J. D., & Grove, P. M. (2013). Segmentation by depth does not always facilitate visual search. *Journal of Vision*, *13*(8):11, 1–14, <https://doi.org/10.1167/13.8.11>. [PubMed] [Article]
- Finlayson, N. J., Zhang, X., & Golomb, J. D. (2017). Differential patterns of 2D location versus depth decoding along the visual hierarchy. *NeuroImage*, *147*, 507–516.
- Fukuda, K., Vogel, E., Mayr, U., & Awh, E. (2010). Quantity, not quality: The relationship between fluid intelligence and working memory capacity. *Psychonomic Bulletin & Review*, *17*(5), 673–679.
- Golomb, J. D. (2015). Divided spatial attention and feature-mixing errors. *Attention, Perception & Psychophysics*, *77*(8), 2562–2569.
- Gorgoraptis, N., Catalao, R. F. G., Bays, P. M., & Husain, M. (2011). Dynamic updating of working memory resources for visual objects. *The Journal of Neuroscience*, *31*(23), 8502–8511.
- Grill-Spector, K., & Malach, R. (2004). The human visual cortex. *Annual Review of Neuroscience*, *27*, 649–677.
- Harrison, S. A., & Tong, F. (2009, April 2). Decoding reveals the contents of visual working memory in early visual areas. *Nature*, *458*(7238), 632–635.
- He, Z. J., & Nakayama, K. (1995). Visual attention to surfaces in three-dimensional space. *Proceedings of the National Academy of Sciences, USA*, *92*(24), 11155–11159.
- Hollingworth, A. (2007). Object-position binding in visual memory for natural scenes and object arrays. *Journal of Experimental Psychology: Human Perception and Performance*, *33*(1), 31–47.
- Hollingworth, A., & Rasmussen, I. P. (2010). Binding objects to locations: The relationship between object files and visual working memory. *Journal of Experimental Psychology: Human Perception and Performance*, *36*(3), 543–564.
- Jiang, Y., Olson, I. R., & Chun, M. M. (2000). Organization of visual short-term memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *26*(3), 683–702.
- Kooi, F. L., Toet, A., Tripathy, S. P., & Levi, D. M. (1994). The effect of similarity and duration on spatial interaction in peripheral vision. *Spatial Vision*, *8*(2), 255–279.
- Lehmkuhle, S., & Fox, R. (1980). Effect of depth separation on metacontrast masking. *Journal of Experimental Psychology: Human Perception and Performance*, *6*(4), 605–621.
- Linke, A. C., Vicente-Grabovetsky, A., Mitchell, D. J., & Cusack, R. (2011). Encoding strategy accounts for individual differences in change detection measures of VSTM. *Neuropsychologia*, *49*(6), 1476–1486.
- Luck, S. J., & Vogel, E. K. (1997, November 20). The capacity of visual working memory for features and conjunctions. *Nature*, *390*(6657), 279–281.
- Ma, W. J., Husain, M., & Bays, P. M. (2014). Changing concepts of working memory. *Nature Neuroscience*, *17*(3), 347–356.

- Marshak, W., & Sekuler, R. (1979, September 28). Mutual repulsion between moving visual targets. *Science*, 205(4413), 1399–1401.
- Maunsell, J. H., & Newsome, W. T. (1987). Visual processing in monkey extrastriate cortex. *Annual Review of Neuroscience*, 10, 363–401.
- McCarley, J. S., & He, Z. J. (2001). Sequential priming of 3-D perceptual organization. *Perception & Psychophysics*, 63(2), 195–208.
- Nakayama, K., He, Z. J., & Shimojo, S. (1995). Visual surface representation: A critical link between lower-level and higher-level vision. *Visual cognition: An invitation to cognitive science* (Vol. 2, pp. 1–70). Cambridge, MA: MIT press.
- Nakayama, K., Shimojo, S., & Silverman, G. H. (1989). Stereoscopic depth: Its relation to image segmentation, grouping, and the recognition of occluded objects. *Perception*, 18(1), 55–68.
- Nakayama, K., & Silverman, G. H. (1986, March 20). Serial and parallel processing of visual feature conjunctions. *Nature*, 320(6059), 264–265.
- Olson, I. R., & Marshuetz, C. (2005). Remembering “what” brings along “where” in visual working memory. *Perception & Psychophysics*, 67(2), 185–194.
- Panichello, M. F., DePasquale, B., Pillow, J. W., & Buschman, T. (2018). Error-correcting dynamics in visual working memory. *bioRxiv*. Retrieved from <https://www.biorxiv.org/content/early/2018/05/10/319103.abstract>
- Papathomas, T. V., Feher, A., Julesz, B., & Zeevi, Y. (1996). Interactions of monocular and cyclopean components and the role of depth in the Ebbinghaus illusion. *Perception*, 25(7), 783–795.
- Pashler, H. (1988). Familiarity and visual change detection. *Perception & Psychophysics*, 44(4), 369–378.
- Pasternak, T., & Greenlee, M. W. (2005). Working memory in primate sensory systems. *Nature Reviews Neuroscience*, 6(2), 97–107.
- Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision*, 10(4), 437–442.
- Pertsov, Y., & Husain, M. (2014). The privileged role of location in visual working memory. *Attention, Perception & Psychophysics*, 76(7), 1914–1924.
- Phillips, W. A. (1974). On the distinction between sensory storage and short-term visual memory. *Perception & Psychophysics*, 16(2), 283–290.
- Postle, B. R., Awh, E., Serences, J. T., Sutterer, D. W., & D’Esposito, M. (2013). The positional-specificity effect reveals a passive-trace contribution to visual short-term memory. *PLoS One*, 8(12), e83483.
- Qian, J., Li, J., Wang, K., Liu, S., & Lei, Q. (2017). Evidence for the effect of depth on visual working memory. *Scientific Reports*, 7(1), 6408.
- Rademaker, R. L., Bloem, I. M., De Weerd, P., & Sack, A. T. (2015). The impact of interference on short-term memory for visual orientation. *Journal of Experimental Psychology: Human Perception and Performance*, 41(6), 1650–1665.
- Rademaker, R. L., Chunharas, C., & Serences, J. T. (2018). Simultaneous representation of sensory and mnemonic information in human visual cortex. *bioRxiv*, <https://doi.org/10.1101/339200>.
- Rademaker, R. L., Park, Y. E., & Sack, A. T. (2018). Evidence of gradual loss of precision for simple features and complex objects in visual working memory. *Journal of Experimental Psychology: Human Perception and Performance*, 44(6), 925–940. Retrieved from <http://psycnet.apa.org/record/2018-08188-001>
- Rademaker, R. L., Tredway, C., & Tong, F. (2012). Introspective judgments predict the precision and likelihood of successful maintenance of visual working memory. *Journal of Vision*, 12(13):21, 1–13, <https://doi.org/10.1167/12.13.21>. [PubMed] [Article]
- Rauber, H. J., & Treue, S. (1998). Reference repulsion when judging the direction of visual motion. *Perception*, 27(4), 393–402.
- Reeves, A., & Lei, Q. (2014). Is visual short-term memory depthful? *Vision Research*, 96, 106–112.
- Schurgin, M. W., Wixted, J. T., & Brady, T. F. (2018). Psychological scaling reveals a single parameter framework for visual working memory. *bioRxiv*. Retrieved from <https://www.biorxiv.org/content/biorxiv/early/2018/05/18/325472.full.pdf>
- Scocchia, L., Cicchini, G. M., & Triesch, J. (2013). What’s “up”? Working memory contents can bias orientation processing. *Vision Research*, 78, 46–55.
- Serences, J. T. (2016). Neural mechanisms of information storage in visual short-term memory. *Vision Research*, 128, 53–67.
- Serences, J. T., Ester, E. F., Vogel, E. K., & Awh, E. (2009). Stimulus-specific delay activity in human primary visual cortex. *Psychological Science*, 20(2), 207–214.
- Sereno, M. I., Dale, A. M., Reppas, J. B., Kwong, K. K., Belliveau, J. W., Brady, T. J., . . . Tootell, R. B. (1995, May 12). Borders of multiple visual areas in humans revealed by functional magnetic resonance imaging. *Science*, 268(5212), 889–893.

- Sereno, M. I., Pitzalis, S., & Martinez, A. (2001, November 9). Mapping of contralateral space in retinotopic coordinates by a parietal cortical area in humans. *Science*, *294*(5545), 1350–1354.
- Shin, H., Zou, Q., & Ma, W. J. (2017). The effects of delay duration on visual working memory for orientation. *Journal of Vision*, *17*(14):10, 1–24, <https://doi.org/10.1167/17.14.10>. [PubMed] [Article]
- Simons, D. J., & Levin, D. T. (1997). Change blindness. *Trends in Cognitive Sciences*, *1*(7), 261–267.
- Snow, J. C., Skiba, R. M., Coleman, T. L., & Berryhill, M. E. (2014). Real-world objects are more memorable than photographs of objects. *Frontiers in Human Neuroscience*, *8*, 837.
- Sprague, T. C., Ester, E. F., & Serences, J. T. (2014). Reconstructions of information in visual spatial working memory degrade with memory load. *Current Biology*, *24*(18), 2174–2180.
- Sreenivasan, K. K., Curtis, C. E., & D’Esposito, M. (2014). Revisiting the role of persistent neural activity during working memory. *Trends in Cognitive Sciences*, *18*(2), 82–89.
- Störmer, V. S., Alvarez, G. A., & Cavanagh, P. (2014). Within-hemifield competition in early visual areas limits the ability to track multiple objects with attention. *The Journal of Neuroscience*, *34*(35), 11526–11533.
- Suchow, J. W., Brady, T. F., Fougner, D., & Alvarez, G. A. (2013). Modeling visual working memory with the MemToolbox. *Journal of Vision*, *13*(10):9, 1–8, <https://doi.org/10.1167/13.10.9>. [PubMed] [Article]
- Talbot, S. A., & Marshall, W. H. (1941). Physiological studies on neural mechanisms of visual localization and discrimination. *American Journal of Ophthalmology*, *24*(11), 1255–1264.
- Theeuwes, J., Atchley, P., & Kramer, A. F. (1998). Attentional control within 3-D space. *Journal of Experimental Psychology: Human Perception and Performance*, *24*(5), 1476–1485.
- Treisman, A., & Zhang, W. (2006). Location and binding in visual working memory. *Memory & Cognition*, *34*(8), 1704–1719.
- 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. *Proceedings of the National Academy of Sciences, USA*, *109*(22), 8780–8785.
- Vogel, E. K., McCollough, A. W., & Machizawa, M. G. (2005, November 24). Neural measures reveal individual differences in controlling access to working memory. *Nature*, *438*(7067), 500–503.
- Wilken, P., & Ma, W. J. (2004). A detection theory account of change detection. *Journal of Vision*, *4*(12):11, 1120–1135, <https://doi.org/10.1167/4.12.11>. [PubMed] [Article]
- Xu, Y., & Nakayama, K. (2007). Visual short-term memory benefit for objects on different 3-D surfaces. *Journal of Experimental Psychology: General*, *136*(4), 653–662.
- Zhang, W., & Luck, S. J. (2008, May 8). Discrete fixed-resolution representations in visual working memory. *Nature*, *453*, 233.