



Working memory for stereoscopic depth is limited and imprecise—evidence from a change detection task

Jiehui Qian¹ · Ke Zhang¹

Published online: 6 August 2019
© The Psychonomic Society, Inc. 2019

Abstract

Most studies on visual working memory (VWM) and spatial working memory (SWM) have employed visual stimuli presented at the fronto-parallel plane and few have involved depth perception. VWM is often considered as a memory buffer for temporarily holding and manipulating visual information that relates to visual features of an object, and SWM for holding and manipulating spatial information that concerns the spatial location of an object. Although previous research has investigated the effect of stereoscopic depth on VWM, the question of how depth positions are stored in working memory has not been systematically investigated, leaving gaps in the existing literature on working memory. Here, we explore working memory for depth by using a change detection task. The memory items were presented at various stereoscopic depth planes perpendicular to the line of sight, with one item per depth plane. Participants were asked to make judgments on whether the depth position of the target (one of the memory items) had changed. The results showed a conservative response bias that observers tended to make ‘no change’ responses when detecting changes in depth. In addition, we found that similar to VWM, the change detection accuracy degraded with the number of memory items presented, but the accuracy was much lower than that reported for VWM, suggesting that the storage for depth information is severely limited and less precise than that for visual information. The detection sensitivity was higher for the nearest and farthest depths and was better when the probe was presented along with the other items originally in the memory array, indicating that how well the to-be-stored depth can be stored in working memory depends on its relation with the other depth positions.

Keywords Working memory · Depth perception · Change detection · Binocular disparity

Introduction

For decades, working memory has been a heated research topic in cognitive science (Baddeley, 2012). Baddeley and Hitch (1974) proposed a multicomponent model for working memory that contains the central executive, the phonological loop, and the visuospatial sketchpad. The visuospatial sketchpad involves temporarily holding (typically at least for 1 s) and manipulating visual information that relates to visual features of an object such as color, shape, or orientation (visual working memory, VWM), and spatial information that concerns the spatial location of an object (spatial working memory, SWM) (Baddeley, 2012). Although the type of information suggested to

be stored in VWM and SWM seems to be independent, the two stores are not necessarily exclusive. For example, VWM is shown to be affected by spatial proximity (Peterson & Berryhill, 2013; Woodman et al., 2003). It has also been suggested that visual features may be linked by location to form object (Treisman & Gelade, 1980), and there is a long-lasting large-capacity, but fragile VWM that bounded by location before a robust VWM can be formed (Sligte et al., 2008).

The robust VWM has been studied extensively (Luck & Vogel, 2013; Ma et al., 2014) and shown to be severely capacity-limited: memory performance deteriorates dramatically when the number of to-be-remembered items exceeds four (Alvarez & Cavanagh, 2004; Awh et al., 2007; Cowan, 2001; Luck & Vogel, 1997; Vogel et al., 2001; Zhang & Luck, 2008). Despite its limited capacity, studies showed that cognitive factors such as selective attention (Astone et al., 2012; Griffin & Nobre, 2003; Heuer & Schubö, 2016; Murray et al., 2013; Qian et al., 2019), Gestalt grouping (Jiang et al., 2004; Li et al., 2018; Xu, 2006), and familiarity (Jackson & Raymond, 2008) could affect VWM performance.

✉ Jiehui Qian
qianjh3@mail.sysu.edu.cn

¹ Department of Psychology, Sun Yat-Sen University, Guangzhou, China

For example, attention can be directed toward certain locations or representations in VWM, improving memory for the relevant items in the focus of attention, and Gestalt grouping could enhance memory performance if visual stimuli are properly chunked. These factors aid VWM to deal with the limited ability of the visual system to process and store the abundant information that is simultaneously presented.

Compared to VWM, it seems that memory capacity for SWM can be much larger. Pioneer work by Vincent Di Lollo showed a large-capacity spatial memory out to 200 ms (Di Lollo, 1977), and several more recent studies showed that spatial information can be held up to 2000 ms (Schneegans & Bays, 2016; Schurgin & Flombaum, 2014; Sheth & Shimojo, 2001). Despite the greater capacity, past research shows that a systematic bias could be observed when reproducing object locations from memory, with the locations attracting or repelling one another (Sheth & Shimojo, 2001; Liverence & Scholl, 2011), or with respect to boundaries and landmarks (Diedrichsen et al., 2004; Huttenlocher et al., 1991; Nelson & Chaiklin, 1980). Since studies on SWM often involved reproducing the location of a briefly presented object by a mouse click or a finger touch (Schneegans & Bays, 2016; Schurgin & Flombaum, 2014), researchers suggest that the biased responses may not result from distorted spatial memory representations, but rather from ‘sampling’ probabilistic memories into discrete actions (Schurgin & Flombaum, 2014). In other words, the memory representation for locations is unbiased but noisy, and distorted responses arise from the coordination between memory and action. However, since Sheth and Shimojo (2001) also used a non-action-based experiment to show the same memory distortion, no decisive conclusion has been reached on whether the effect is due to sensorimotor factors.

Most studies on VWM and SWM presented visual stimuli on a two-dimensional (2-D) fronto-parallel plane with no depth information involved. Recently, several studies investigated the effect of depth perception on VWM by presenting the to-be-remembered items at various stereoscopic depth planes (Chunharas et al., 2019; Qian et al., 2017; Qian et al., 2018; Reeves & Lei, 2014; Xu & Nakayama, 2007). Although Reeves and Lei (2014) did not find clear evidence that VWM can be improved through separating items by stereoscopic depth using a partial report paradigm, Xu and Nakayama (2007) and Chunharas et al. (2019) found a beneficial effect of depth on VWM and Qian et al. (2017, 2018) showed that memory performance could be affected by depth location of an item using an adapted change-detection paradigm. However, these studies all focused on whether the storage for visual information can be affected by depth information, not how depth

information per se is stored in working memory. There is one study by Reeves and Lei (2017) that investigated working memory for numerals that associated with depths using a partial report paradigm. In their study, numerals were shown at different depth planes, followed after various delays by an arrow cue to indicate one of the depth planes, and the participants needed to report the numeral whose depth was cued by the arrow. The study showed that the memory performance first fell with increasing cue delay but then recovered gradually, and the authors suggested that information concerning the depths of the numerals first decays in sensory memory but then transfers to a visual memory specialized for depth. Although this study tackled the question of storage for depth information, their investigation on memory for depth is embedded in investigating memory for numerals, which renders their conclusion on memory for depth to critically rely on the strength of association between a numeral and its depth position. Therefore, we think that the question of how depth positions are stored in working memory has not been explicitly investigated, leaving gaps in the existing literature on working memory.

Since studies suggest that the visual system has separate processing mechanisms for depth information and 2-D (fronto-parallel) spatial information (Umemura, 2015; Finlayson & Golomb, 2016; Finlayson et al., 2017), it is possible that working memory for depth also has characteristics distinct from what has been reported for VWM and SWM for 2-D spatial locations. Investigating on this topic may provide us insights on the underlying mechanisms of working memory for depth information. Here we explored working memory for depth information by using a change detection paradigm. The change detection task (CDT), which has been frequently used in studies on VWM (Luck & Vogel 1997, 2013), has been proven to be a valid tool for investigating working memory and has been used by early psychologists to test spatial memory (Di Lollo, 1977; Peterson et al., 1977). In a CDT, observers were instructed to detect any change between a briefly presented memory array and a probe after a period of retention. In our study, multiple memory items were simultaneous presented with each occupying a different depth plane, and participants were required to remember the positions in depth of these items. Employing a CDT allows us to make comparisons between the change detection accuracy for depth position and that reported for VWM. By manipulating the number of memory items presented and whether or not the probe was presented alone or with the other memory items, we investigated how these factors could affect the memory performance for depth.

Method

Participants

Thirty-two participants (three males; mean age = 22.6 years) from Sun Yat-Sen University (SYSU) with normal or corrected-to-normal vision took part in the experiment. They were required to pass a screening test to ensure that they could accurately perceive the disparity-defined depth. Eleven out of 32 participants that passed the test were recruited for the experiment (ten for the formal experiment and five for the control experiment) and received payment for their participation. This research was approved by the SYSU Institutional Review Board (IRB). Written informed consent approved by the IRB was obtained from each participant prior to all of the experiments.

Stimuli

The stimuli were viewed against a uniform gray background (102 cd/m^2) through a Wheatstone stereoscope on a pair of 21-inch ViewSonic monitors. The display resolution was set to 1920×1080 pixels, with a refresh rate of 60 Hz. For the typical viewing distance of 75 cm, a pixel subtended approximately 1 arcmin.

The memory array was composed of a set of blue squares (25.8 cd/m^2) that were arranged in a circular configuration with a radius of 3.5° from the center of the screen (see Fig. 1). The set sizes of the memory array were 1, 2, 4, and 6. Any two nearest neighboring items were separated by 3.5° . The memory items were presented at various depth planes perpendicular to the line of sight, with one item per depth plane. The depth position of a memory item was randomly selected from a set of seven depth planes without replacement. The depth planes were separated by relative disparities ranging from $-.51^\circ$ to $.51^\circ$ with a step of $.17^\circ$, which corresponds to -7.0, -4.8, -2.5, 0, 2.7, 5.5, and 8.6 cm from the monitor screen with a typical inter-pupillary distance of 6.5 cm. These disparities were selected so that the left- and right-eye images could be reliably fused and the items clearly appeared to be separated in depth (Blakemore, 1970). Each item subtended approximately $.80^\circ \times .80^\circ$ of visual angle. Due to the size constancy phenomenon, the items at far planes appeared to be larger than those at the near planes. We measured the perceived size at each depth plane for four participants using a nulling paradigm and the results showed that on average the perceived size increased by 1% as its depth plane receded to the next farther depth plane. Therefore, we used the item with zero-disparity as a standard and scaled the size of the items at the other six

planes accordingly so that the size of the items at all planes appeared to be the same.

Procedure

Observers were seated in a dark room to complete the experiment. First they needed to pass a screening task to ensure that they could well perceive the stereoscopic depth. In each trial, two horizontally displaced blue squares were presented for 500 ms. The depth position of one item was randomly selected from the seven depth planes, and the other item was separated by a relative disparity of $.17^\circ$. Participants were instructed to judge which item was farther as quickly as possible, and were required to achieve an accuracy of above 90% for 48 trials in order to continue with the formal experiment.

After the participants passed the screening test, they were trained for a short time (2–5 min) to get acquainted with the stimuli and the task in the formal experiment. Each trial began with a fusion phase that a red cross subtended $.65^\circ \times .65^\circ$ was presented at the center of the screen. The participant was instructed to fixate at the cross and fuse the left-eye and right-eye images of the red cross until no double image was perceived. He or she then confirmed the success of fusion by pressing a key, and the red cross turned black and persisted throughout the trial. After the fixation cross turned black for 400 ms, the memory array that composed of blue squares was presented for 800 ms. It was followed by a 900-ms retention interval and then the test phase. Two types of probe display were tested. In the *single-display* block, a probe was shown at one of the frontal-parallel locations (test location) in the memory array; in the *whole-display* block, a probe that was indicated by a black frame ($1.3^\circ \times 1.3^\circ$) was shown along with the other memory items. Note that with a set size of 1, a single display was equivalent to a whole display; therefore only set sizes of 2, 4, and 6 were tested for the whole display. The depth position of the probe would either remain the same as in the memory array or be randomly selected from the depth planes that had not been previously chosen in the memory array. The test item remained on the screen until the observer responded. After the response, a 1000-ms blank intertrial interval was presented before the next trial. A diagram of the task sequence is shown in Fig. 1.

The participants were asked to judge whether the depth position of the probe had changed. They needed to press “1” on the keyboard if ‘a change’ was perceived and press “3” if ‘no change’ was perceived. On 50% of the trials, the depth position of the probe changed. There was an equal number of trials for each set size and for each depth plane

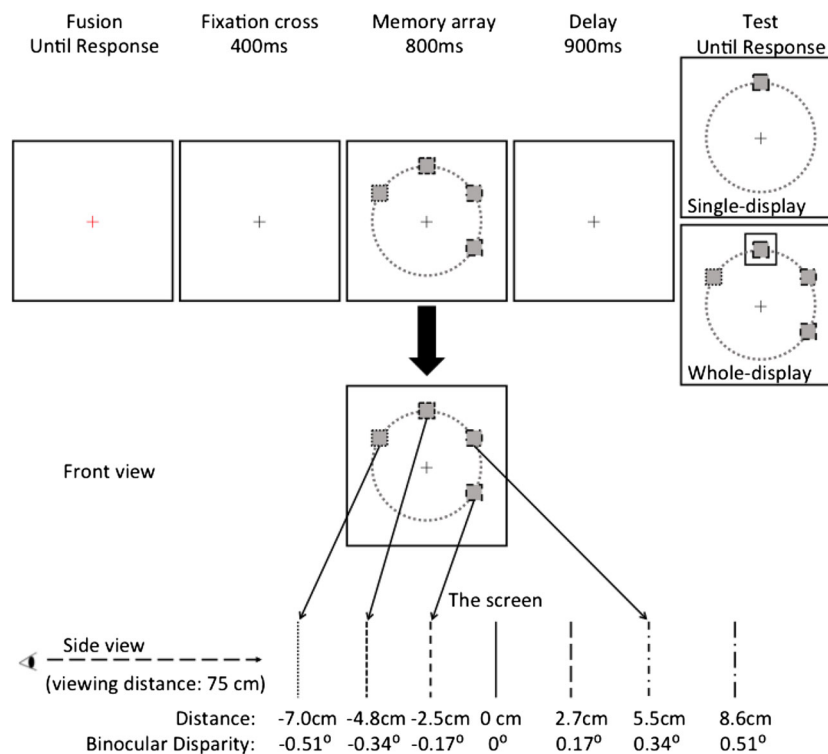


Fig. 1 Stimuli and procedure in the experiment. *Top*: task sequence. *Bottom*: the front view and the side view of the memory display. The memory items were arranged in a circular configuration (indicated by a *dashed circle*), and were outlined with various types of lines to indicate their different depth positions. No line or circle was presented in the actual experiment

that was selected to be tested (target depth). Each observer completed a total of 1120 trials for each display block, with 40 trials per target depth per set size. Half of the observers completed the single-display block first, and the other half completed the whole-display block first. The order of the trials was randomized during the experiment.

In addition, a control experiment examining the perceptual performance on depth was carried out to test whether observers could successfully detect changes in depth under the condition of no retention with the current experimental setting. The stimuli and procedure were identical to that in the formal experiment, except for the following changes: 1) there was no delay between the memory display and the probe display; 2) a set size of 4 or 6 was used as a preliminary testing showed that the change detection accuracy approached 100% for set size of 1 or 2; 3) there were only ‘change’ trials (equal chance on a ‘forward’ change and a ‘backward’ change) and the observers were asked to report whether the depth position of the probe was nearer or farther than the target depth. We employed this task since a preliminary testing showed that the change detection accuracy approached 100% if the task was to detect ‘whether there is a change’ on 50% of trials as in the formal experiment. Each

observer completed a total of 280 trials for each display block, with 20 trials per target depth per set size.

Data analysis

A $2 \times 3 \times 7$ (display \times set size \times target depth) repeated-measures ANOVA was conducted on the change detection performance. Because a set size of 1 was only used in the single display, this set of data was not included in the analysis. The performance was calculated in terms of signal detection theory, in which the accuracy or detection sensitivity (A') and the response bias (Grier's B'') were assessed separately. A' (Pollack & Norman, 1964; Stanislaw & Todorov, 1999) was used in our study instead of d' , since some researchers suggested that it was more accurate than d' (Donaldson, 1993). Grier's B'' was used instead of β since B'' does not have the indeterminacy of β when an observer makes no false alarm responses. B'' ranges from -1 to 1. A positive value indicates a conservative bias (in favor of ‘no change’ responses), and a negative value indicates a liberal bias (in favor of ‘change’ responses).

The performance in the control experiment was calculated in terms of change detection accuracy (percent of

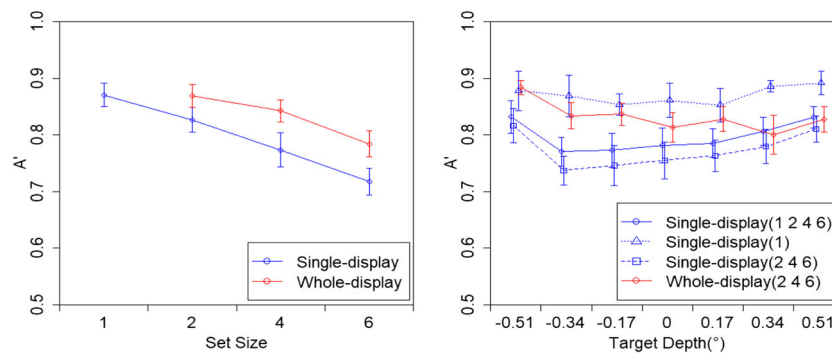


Fig. 2 Results of A' as a function of set size (*left panel*), and target depth (*right panel*). The A' 's were averaged across target depths and set sizes (indicated by the *numbers in the parentheses*), respectively, since there was no significant interaction effect

correct), and was analyzed using a $2 \times 2 \times 7$ (display \times set size \times target depth) repeated-measures ANOVA.

Results

The mean A' was about .88 for a set size of 1 (single display). The repeated-measures ANOVA on A' showed a significant main effect of display, [$F(1, 9) = 5.09, p = 0.05; \eta_p^2 = .36$]. The mean A' was higher for the whole display (.83) than for the single display (.77). The main effect of set size was significant, [$F(2, 18) = 45.23, p < 0.001; \eta_p^2 = .83$]. The mean A' decreased with set size for the two displays, and the trend analysis revealed a significant linear trend, [$F(1, 9) = 53.71, p < 0.001; \eta_p^2 = .86$]. The main effect of target depth was significant, [$F(6, 54) = 3.38, p = 0.007; \eta_p^2 = .27$]. There was a significant quadratic trend that A' first decreased and then increased with target depth, [$F(1, 9) = 6.48, p = 0.031; \eta_p^2 = .42$]. There were no significant two-way or three-way interactions found, p values > 0.05 . In Fig. 2,

the left and right panels show the mean A' as a function of set size and target depth, respectively.

Figure 3 shows the mean Grier's B'' as a function of set size (Fig. 3, the left panel) and target depth (Fig. 3, the right panel). The mean B'' was positive at each set size level and each target depth, indicating that participants tended to make 'no change' responses. It was about .68 for a set size of 1 (single display). The repeated-measures ANOVA on Grier's B'' showed a significant main effect of set size, [$F(2, 18) = 1.31, p = 0.001; \eta_p^2 = .54$]. The mean B'' decreased with set size for the two displays, showing a significant linear trend, [$F(1, 9) = 11.75, p = 0.008; \eta_p^2 = .56$]. The main effect of display, [$F(1, 9) = 1.08, p = 0.324; \eta_p^2 = .11$], or target depth, [$F(3, 27) = 1.36, p = 0.246; \eta_p^2 = .13$], was not significant. The interaction of display and target depth was significant, [$F(6, 54) = 3.12, p = 0.011; \eta_p^2 = .26$]. Simple effect analysis showed that the effect of target depth was not significant for the single display, [$F(6, 54) = 1.69, p = 0.14; \eta_p^2 = .16$], but was significant for the whole display (a quadratic trend), [$F(6, 54) = 2.41, p = 0.039; \eta_p^2 = .21$]. In other words,

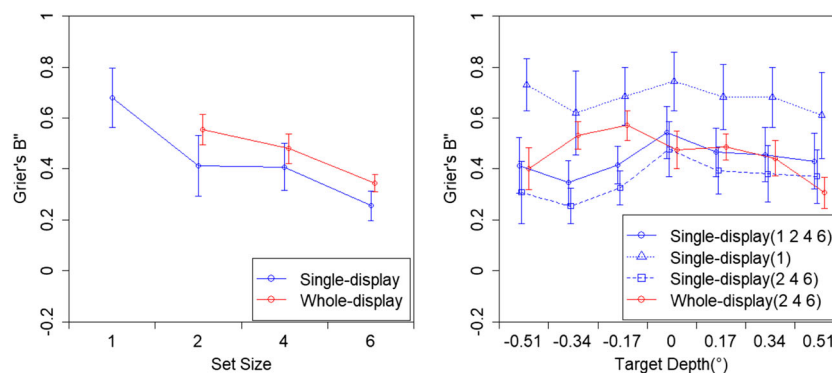


Fig. 3 Results of Grier's B'' as a function of set size (*left panel*), and target depth (*right panel*). The B'' 's were averaged across target depths and set sizes (indicated by the *numbers in the parentheses*), respectively

Table 1 Accuracy (std. err.) for different set sizes, target depths, and display types in the control experiment

Target depth	Single display		Whole display	
	Set size 4	Set size 6	Set size 4	Set size 6
-.51°	.94 (.04)	.96 (.02)	.96 (.02)	.98 (.02)
-.34°	1.00 (.00)	.96 (.02)	1.00 (.00)	.98 (.02)
-.17°	1.00 (.00)	.96 (.03)	1.00 (.00)	.98 (.02)
0°	.94 (.04)	.94 (.03)	.98 (.02)	1.00 (.00)
.17°	.96 (.02)	.90 (.06)	.98 (.02)	.96 (.04)
.34°	.94 (.02)	.88 (.06)	.94 (.02)	.94 (.04)
.51°	.90 (.11)	.92 (.08)	.90 (.06)	.88 (.06)

the mean of B' first increased and then decreased with target depth in the whole display. No other interaction was found, p values > 0.5.

For the control experiment, the mean accuracies were high for each set size and target depth regardless of display (see Table 1). The repeated-measures ANOVA on accuracy showed no significant main effect of display [$F(1, 4) = 5.57$, $p = 0.078$; $\eta_p^2 = .28$], set size [$F(1, 4) = 5.00$, $p = 0.089$; $\eta_p^2 = .26$], and target depth [$F(6, 24) = 5.25$, $p = 0.058$; $\eta_p^2 = .36$]. No significant interaction was found, p values > 0.324.

Discussion

The present study investigated working memory for depth using a change detection task. We found that the detection sensitivity decreased with the number of memory items presented (set size), and it was enhanced when presenting the probe together with the other memory items (whole display) compared to when presenting the probe alone (single display). The sensitivity was higher for detecting changes in the nearest depth plane regardless of the probe display. Overall, observers tended to make conservative responses when detecting changes in depth, and the response bias varied with the depth position to be held in memory (target depth) in the whole display. Because the observed memory performance for depth information seems to be completely different from what has been reported for VWM or SWM in past literature (see discussions below), we propose that there might be a separate cognitive buffer (depth working memory, DWM) for temporary storage and manipulation of depth information, which may add as a new building block to the existing architecture of the multicomponent model of working memory.

The change detection accuracy for DWM was much lower than that reported for VWM. Compared to a nearly 100% of correct for memorizing one feature or object (Luck & Vogel, 1997), the percent of correct in this experiment

was about 78% with a set size of 1. This is consistent with that reported by Reeves and Lei (2017), who also demonstrated the poor retention of numerals associated with depths even for a set size of 1 using a partial report paradigm. In addition, we compared the performance from four observers who also participated in a change detection task for color, and found that their accuracy was about 97% for VWM and 73% for DWM with a set size of 4, and was about 83% for VWM and 69% for DWM with a set size of 6. The poor memory performance for depth cannot be due to the imprecision of stereopsis (McKee et al., 1990), since all observers achieved an accuracy of 90% or above in the screening task where they successfully discriminated two briefly presented depth planes, and the perceptual task in the control experiment showed that the accuracy was high even for a set size of 4 or 6. These results suggest that although we are adequately capable of perceiving changes in stereoscopic depths, our ability to retain a representation of metric depth in working memory is severely limited, indicating that our memory for depth position is probably imprecise.

Despite the overall unsatisfactory memory performance, the accuracy for the nearest target depth was the highest regardless of the probe display. This 'front advantage' is in accordance with previous studies, where researchers found that performance was better for perceptually closer objects in visual search tasks (Ogawa & Macaluso, 2015) and memory tasks (Qian et al. 2017, 2018). It is suggested that objects near us may be more ecologically useful and important, therefore we may prioritize the near objects for processing and encoding compared with the distant ones (Qian et al., 2017). Similarly, near depth positions may also be prioritized for processing and storage because of its ecological relevancy. On the other hand, the performance for the farthest depth was also better than the intermediate ones, especially for the single display. This benefit for the farthest depth is only found in DWM, as previous studies show that VWM performance for objects located at a farther depth was worse than for those located at a nearer

depth (Qian et al. 2017, 2018). In general, the improved performance on the boundary depth positions is analogous to the primacy and recency effects typically found in free-call tasks for words displayed serially (Postman & Phillips, 1965), which are suggested to reflect contributions from long-term memory and short-term memory, respectively. However, this explanation does not apply to the memory advantage for boundary positions, and we suggest that this memory advantage may relate to the mechanism of how depth positions are perceived and encoded into memory. Since research shows that the presence of a more distant object improves the judgments on an object's position in depth (Blank, 1958; Foley, 1985; Gogel, 1972; Sousa et al., 2011), it is possible that the nearest and farthest depths are better perceived than the intermediate depths, and perception of boundaries helps to enhance the precision of perceiving the intermediate depths. In other words, the boundary depth positions may be prioritized in encoding and processing, resulting a stronger 'input strength', and therefore are better stored in working memory.

Furthermore, the memory performance for the whole display was improved compared to the single display. In a whole display, the probe was presented with the other memory items, whose depth positions remained the same throughout the trial and therefore might serve as references to help the observers estimate whether the relative depth separation had changed among items. Research shows that depth perception could be enhanced by providing references (Andrews et al., 2001; Petrov & Glennerster, 2004; Petrov et al., 2006), therefore it is possible that DWM may as well be improved when reliable references were available. Conversely, it is suggested that unreliable references could deteriorate the memory performance for spatial locations. For example, in a memory-guided reaching task, Fiehler et al. (2007) found that varying the positions of local references could systematically bias the accuracy for reproducing the spatial location of an object.

The analysis on response criteria shows a reliable bias that participants tend to make conservative responses when detecting changes in depth. Since the high accuracy in the control experiment showed that stereoscopic depth was well perceived and changes in depth could be detected, the conservative bias may result from rapid loss of precise depth information stored in DWM and consequently a failure to discriminate the probe's depth position from the stored target depth, with the 'no change' response as a default. The conservative bias decreases with set size, and is slightly larger for the intermediate-depth planes in a whole display, i.e., the observers are less likely to report that a to-be-remembered depth has changed when it is originally located around the middle of the tested depth range. One of the

possible explanations is that the potential magnitude of change for the intermediate depths is smaller than that for the boundary depths, since the depth position of the probe was randomly chosen from the 'pool' of seven depth planes on the 'change' trials. Presumably, a small magnitude of change can be harder to detect, and this may aggravate the conservative bias for the intermediate depth planes. However, we think that this possibility is unlikely based on the following reasons: 1) it cannot explain why this bias pattern is not observed for the single display; 2) our control experiment with an identical setting in ranges of change magnitudes showed that observers could well detect changes in depth regardless of target depth. In other words, the results of the perceptual task indicate that even the smallest magnitude of change (.17°) in our experiment was large enough to reach the response criterion for a 'different' judgment. Another possibility is that this bias pattern arises from a comparison process during the test phase. Compared to a nearest or farthest depth that one can easily 'label', it is harder to define the relation of an intermediate depth among the others. During the test phase, by comparing the relation among different depths, it is more likely for a change in the nearest/farthest depth to lose its 'boundary label' than for a change in the intermediate depth to obtain a 'boundary label'. In other words, 'boundary label' switches more often for changes in the nearest and farthest depth than for changes in an intermediate depth. Therefore one may be less certain about whether the intermediate depth has changed or not and be more reluctant to make a confirmative response that it has changed. A whole display may allow and encourage the comparison strategy more than a single display does, so the bias pattern is not shown in the latter.

Finally, here we used A' to estimate the detection sensitivity for depth, but note that there might be concerns regarding this measure. Pastore et al. (2003) showed that A' is equivalent to d' regardless of response bias when sensitivity is low, but tends to be underestimated when sensitivity is high and response bias is not neutral. Although we found a conservative bias when detecting changes in depth, the detection sensitivity was low ($mean A' = .81$). Based on the isosensitivity contour of $A' = 0.8$ provided by Pastore et al. (2003), there is only a small amount of underestimation for A' when bias is extreme. Therefore, we think that the influence of a possible underestimate for A' is negligible in our study.

Conclusions

Working memory for depth is rarely investigated in the literature and here we show that the memory performance for detecting changes in stereoscopic depth is generally unsatisfactory but is better for nearest and farthest depths.

This may reflect a ‘boundary advantage’ in DWM. The effect of reference on accuracy and the pattern of the conservative bias indicate that how well a depth position can be stored in working memory depends on its relation with the other depth positions. We conclude that the storage for depth information is limited and less precise than that for visual information and suggest that the relational information between the to-be-stored depths can be crucial in DWM.

Acknowledgements This work has been supported by the National Natural Science Foundation of China (31500919). The authors have no competing financial interests that might be perceived to influence the results and/or discussion reported in this paper.

Open practices statement The data and materials for all experiments are available upon request and none of the experiments was preregistered.

References

- Alvarez, G. A., & Cavanagh, P. (2004). The capacity of visual short-term memory is set both by visual information load and by number of objects. *Psychological Science*, *15*(2), 106–111.
- Andrews, T. J., Glennerster, A., & Parker, A. J. (2001). Stereoacuity thresholds in the presence of a reference surface. *Vision Research*, *41*, 3051–3061.
- Astle, D. E., Summerfield, J., Griffin, I., & Nobre, A. C. (2012). Orienting attention to locations in mental representations. *Attention, Perception, & Psychophysics*, *74*(1), 146–162.
- Awh, E., Barton, B., & Vogel, E. K. (2007). Visual working memory represents a fixed number of items regardless of complexity. *Psychological Science*, *18*(7), 622–628.
- Baddeley, A. (2012). Working memory: theories, models, and controversies. *Annual Review of Psychology*, *63*, 1–29.
- Baddeley, A., & Hitch, G. J. (1974). Working memory. In Bower, G. H. (Ed.) *The psychology of learning and motivation* (Vol. 8). New York: Academic Press.
- Blakemore, C. (1970). The range and scope of binocular depth discrimination in man. *Journal of Physiology*, *211*(3), 599–622.
- Blank, A. A. (1958). Analysis of experiments in binocular space perception. *Journal of the Optical Society of America*, *48*, 911–925.
- 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*(4), 1–16.
- Cowan, N. (2001). Metatheory of storage capacity limits. *Behavioral and Brain Sciences*, *24*(01), 154–176.
- Diedrichsen, J., Werner, S., Schmidt, T., & Trommershäuser, J. (2004). Immediate spatial distortions of pointing movements induced by visual landmarks. *Perception & Psychophysics*, *66*, 89–103.
- Di Lollo, V. (1977). Temporal characteristics of iconic memory. *Nature*, *267*(5608), 241–243.
- Donaldson, W. (1993). Accuracy of d' and a' as estimates of sensitivity. *Bulletin of the Psychonomic Society*, *31*(4), 271–274.
- Fiehler, K., Burke, M., Engel, A., Bien, S., & Rösler, F. (2007). Kinesthetic working memory and action control within the dorsal stream. *Cerebral Cortex*, *18*(2), 243–253.
- 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., Zhang, X., & Golomb, J. D. (2017). Differential patterns of 2D location versus depth decoding along the visual hierarchy. *NeuroImage*, *147*, 507–516.
- Foley, J. M. (1985). Binocular distance perception: Egocentric distance tasks. *Journal of Experimental Psychology: Human Perception and Performance*, *11*, 133–149.
- Gogel, W. C. (1972). Depth adjacency and cue effectiveness. *Journal of Experimental Psychology*, *92*(2), 176.
- Griffin, I. C., & Nobre, A. C. (2003). Orienting attention to locations in internal representations. *Journal of Cognitive Neuroscience*, *15*(8), 1176–1194.
- Heuer, A., & Schubö, A. (2016). Feature-based and spatial attentional selection in visual working memory. *Memory & Cognition*, *44*(4), 621–632.
- Huttenlocher, J., Hedges, L. V., & Duncan, S. (1991). Categories and particulars: Prototype effects in estimating spatial location. *Psychological Review*, *98*, 352–376.
- Jackson, M. C., & Raymond, J. E. (2008). Familiarity enhances visual working memory for faces. *Journal of Experimental Psychology: Human Perception and Performance*, *34*(3), 556.
- Jiang, Y., Chun, M. M., & Olson, I. R. (2004). Perceptual grouping in change detection. *Perception & Psychophysics*, *66*(3), 446–453.
- Li, J., Qian, J., & Liang, F. (2018). Evidence for the beneficial effect of perceptual grouping on visual working memory: an empirical study on illusory contour and a meta-analytic study. *Scientific Reports*, *8*(1), 13864.
- Liverence, B. M., & Scholl, B. J. (2011). Selective attention warps spatial representation: Parallel but opposing effects on attended versus inhibited objects. *Psychological Science*, *22*, 1600–1608.
- 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 in Cognitive Sciences*, *17*(8), 391–400.
- Ma, W. J., Husain, M., & Bays, P. M. (2014). Changing concepts of working memory. *Nature Neuroscience*, *17*(3), 347.
- McKee, S. P., Welch, L., Taylor, D. G., & Bowne, S. F. (1990). Finding the common bond: stereoacuity and the other hyperacuties. *Vision Research*, *30*, 879–891.
- Murray, A. M., Nobre, A. C., Clark, I. A., Cravo, A. M., & Stokes, M. G. (2013). Attention restores discrete items to visual short-term memory. *Psychological Science*, *24*(4), 550–556.
- Nelson, T. O., & Chaiklin, S. (1980). Immediate memory for spatial location. *Journal of Experimental Psychology: Human Learning and Memory*, *6*, 529–545.
- Ogawa, A., & Macaluso, E. (2015). Orienting of visuo-spatial attention in complex 3D space: Search and detection. *Human Brain Mapping*, *36*(6), 2231–2247.
- Pastore, R. E., Crawley, E. J., Berens, M. S., & Skelly, M. A. (2003). “Nonparametric” a' and other modern misconceptions about signal detection theory. *Psychonomic Bulletin & Review*, *10*(3), 556–569.
- Peterson, D. J., & Berryhill, M. E. (2013). The gestalt principle of similarity benefits visual working memory. *Psychonomic Bulletin & Review*, *20*(6), 1282–1289.
- Peterson, L. R., Rawlings, L., & Cohen, C. (1977). The internal construction of spatial patterns. In: *Psychology of learning and motivation* (Vol. 11, pp. 245–276). Elsevier.
- Petrov, Y., & Glennerster, A. (2004). The role of a local reference in stereoscopic detection of depth relief. *Vision Research*, *44*(4), 367–76.

- Petrov, Y., Verghese, P., & McKee, S. P. (2006). Collinear facilitation is largely uncertainty reduction. *Journal of Vision*, *6*(2), 170–8.
- Pollack, I., & Norman, D. A. (1964). A nonparametric analysis of recognition experiments. *Psychonomic Science*, *1*, 125–126.
- Postman, L., & Phillips, L. W. (1965). Short-term temporal changes in free recall. *Quarterly Journal of Experimental Psychology*, *17*(2), 132–138.
- 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. <https://doi.org/10.1038/s41598-017-06719-6>.
- Qian, J., Zhang, K., Lei, Q., Han, Y., & Li, W. (2019). Task-dependent effects of voluntary space-based and involuntary feature-based attention on visual working memory. *Psychological Research Psychologische Forschung*, 1–16. Online Published.
- Qian, J., Zhang, K., Wang, K., Li, J., & Lei, Q. (2018). Saturation and brightness modulate the effect of depth on visual working memory. *Journal of Vision*, *18*(9), 16, 1–12.
- Reeves, A., & Lei, Q. (2014). Is visual short-term memory depthful?. *Vision Research*, *96*, 106–112.
- Reeves, A., & Lei, Q. (2017). Short-term visual memory for location in depth: a u-shaped function of time. *Attention Perception & Psychophysics*, *79*(7), 1917–1932.
- Schneegans, S., & Bays, P. M. (2016). No fixed item limit in visuospatial working memory. *Cortex*, *83*, 181–193.
- Schurgin, M. W., & Flombaum, J. I. (2014). How undistorted spatial memories can produce distorted responses. *Attention Perception & Psychophysics*, *76*, 1371–1380.
- Sheth, B. R., & Shimojo, S. (2001). Compression of space in visual memory. *Vision Research*, *41*(3), 329–341.
- Sligte, I. G., Scholte, H. S., & Lamme, V. A. (2008). Are there multiple visual short-term memory stores? *Plos One*, *3*(2), e1699.
- Sousa, R., Brenner, E., & Smeets, J. B. (2011). Objects can be localized at positions that are inconsistent with the relative disparity between them. *Journal of Vision*, *11*(2), 1–18.
- Stanislaw, H., & Todorov, N. (1999). Calculation of signal detection theory measures. *Behavior Research Methods, Instruments, Computers*, *31*(1), 137–149.
- Treisman, A. M., & Gelade, G. (1980). A feature-integration theory of attention. *Cognitive Psychology*, *12*(1), 97–136.
- Umamura, H. (2015). Independent effects of 2-D and 3-D locations of stimuli in a 3-D display on response speed in a Simon task. *Frontiers in Psychology* 6.
- Vogel, E. K., Woodman, G. F., & Luck, S. J. (2001). Storage of features, conjunctions, and objects in visual working memory. *Journal of Experimental Psychology: Human Perception and Performance*, *27*(1), 92.
- Woodman, G. F., Vecera, S. P., & Luck, S. J. (2003). Perceptual organization influences visual working memory. *Psychonomic Bulletin & Review*, *10*(1), 80–87.
- Xu, Y. (2006). Understanding the object benefit in visual short-term memory: The roles of feature proximity and connectedness. *Perception & Psychophysics*, *68*(5), 815–828.
- 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.
- Zhang, W., & Luck, S. J. (2008). Discrete fixed-resolution representations in visual working memory. *Nature*, *453*(7192), 233–235.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.