Papenmeier, F., & Huff, M. (in press). Viewpoint-dependent representation of contextual information in visual working memory. *Attention, Perception, & Psychophysics*.

The final publication is available at springerlink.com

Viewpoint-dependent representation of contextual information in visual working memory

Frank Papenmeier & Markus Huff

University of Tübingen, Germany

Author Note

Frank Papenmeier, Department of Psychology, University of Tübingen, Germany; Markus Huff, Department of Psychology, University of Tübingen, Germany.

Correspondence concerning this article should be addressed to Frank Papenmeier, University of Tübingen, Schleichstr. 4, D-72076 Tübingen, Germany. E-mail: frank.papenmeier@uni-tuebingen.de

Word count: 3000 (Abstract: 196)

Abstract

Objects are not represented individually in visual working memory (VWM), but in relation to the contextual information provided by other memorized objects. We studied whether the contextual information provided by the spatial configuration of all memorized objects is viewpoint-dependent. We ran two experiments asking participants to detect changes in locations between memory and probe for one object highlighted in the probe image. We manipulated the changes in viewpoint between memory and probe (Experiment 1: 0°, 30°, 60°; Experiment 2: 0°, 60°) as well as the spatial configuration visible in the probe image (Experiment 1: full configuration, partial configuration; Experiment 2: full configuration. no configuration). Location change detection was higher with the full spatial configuration than the partial or no spatial configuration at viewpoint changes of 0°, thus replicating previous findings on the non-independent representation of individual objects in VWM. Most importantly, the effect of spatial configurations decreased with increasing viewpoint changes, suggesting a viewpoint-dependent representation of contextual information in VWM. We discuss these findings within the context of this special issue, in particular whether research performed within the slot-versus-resources debate and research on effects of contextual information might focus on two different storage systems within VWM.

Keywords: visual working memory, contextual information, viewpoint-dependence, spatial configurations

Relocating a letter on the desk is easier if the context surrounding it remains the same than when it has changed. But how is the context represented? Will the context still support relocation, for example, when one approaches the desk from another viewpoint? In the present experiments, we investigate this question with a location change detection paradigm, thereby studying the representation of contextual information, in particular spatial configurations, in visual working memory (VWM).

When memorizing multiple objects in VWM, these objects are not represented independently from one another. Instead, research showed that the configuration formed by the objects contributes to the formation of memory representations. For example, when participants detect changes to a single object highlighted during a test phase, detection accuracy is increased if the original contextual information that was present during the memory phase is also shown (Hollingworth, 2007; Jiang, Chun, & Olson, 2004; Jiang, Olson, & Chun, 2000; Papenmeier, Huff, & Schwan, 2012). Furthermore, effects of context have also been found in studies showing that the memory for the size of an individual object is biased towards the mean size of the memorized objects (Brady & Alvarez, 2011). Based on these findings, it has been suggested that VWM is organized in a hierarchical manner (Brady & Alvarez, 2011; Jiang et al., 2000), representing and integrating information at different levels of abstraction, such as at an item level and at a more global (configurational) level.

Despite the evidence for the role of contextual information for VWM representations, context effects are restricted under some conditions. When participants detect changes to a single object highlighted during the test phase, for example, detection accuracy increases only with the presence of the full spatial configuration from the memory phase, but not with the presence of a partial spatial configuration (Jiang et al., 2000; Papenmeier et al., 2012). This is true even when the partial configuration contains four out of six objects that were present during memory encoding (Papenmeier et al., 2012). Further restrictions have been identified for effects of spatial configurations on color change detection (Jiang et al., 2000), which only occur when the presence of a spatial cueing box implicitly encourages spatial coding (Woodman, Vogel, & Luck, 2012). With the present experiments, we investigate another possible restriction of spatial configuration effects on location change-detection, namely whether they are viewpoint-dependent.

Although effects of viewpoints on the representation of spatial configurations in VWM have not been studied yet, viewpoint-effects are common in many research areas (e.g., Bülthoff & Edelman, 1992; Huff, Papenmeier, Jahn, & Hesse, 2010; Simons & Wang, 1998), and there is related work within the contextual-cueing paradigm. In the contextual-cueing paradigm (Chun & Jiang, 1998), participants perform a visual search and the spatial configuration of some trials is repeated across the experiment. Participants implicitly learn the spatial configuration of these trials, resulting in faster reaction times for trials with repeated configurations than new configurations towards the end of the experiment. This benefit of repeated context is viewpoint-dependent (Chua & Chun, 2003; Jiang, Swallow, & Capistrano, 2013). That is, spatial configurations are only learned implicitly if they are seen from the same viewpoint repeatedly (Jiang et al., 2013) and transfer of learned contextual information to new viewpoints decreases with increasing angles of viewpoint deviation (Chua & Chun, 2003). Therefore, effects of spatial configurations on location change detection performance might also be viewpoint-dependent, suggesting a viewpoint-dependent representation of spatial configurations in VWM.

The prediction of a viewpoint-dependent representation of spatial configurations in VWM is in line with a recent model of VWM (Wood, 2011) proposing at least three independent visual storage systems in VWM: a spatiotemporal storage, an object identity storage, and a snapshot storage. Within this framework, spatial configurations are most likely represented within the snapshot storage. Because the snapshot storage is characterized as storing information in a view-dependent format (Wood, 2011), effects of spatial configuration should decrease with increasing deviations between memory and test viewpoints.

Summarizing, previous research suggests that spatial configurations are stored in a viewpoint-dependent manner in VWM. With the present experiments, we investigated this with a location change detection paradigm with varying contextual information (Jiang et al., 2000; Papenmeier et al., 2012), that we extended to a three-dimensional case allowing for the manipulation of viewpoints through scene rotations.

Experiment 1

Method

Participants

Thirty students participated in exchange for monetary compensation or course credit. All participants reported normal or corrected-to-normal vision.

Apparatus and Stimuli

We presented the stimuli on a 15.4" HP EliteBook 8530p with an ATI Mobility Radeon HD 3650 graphics card at an unrestricted viewing distance of 60 cm. Stimuli were generated using the Blender Game Engine and Python. Each stimulus depicted a scene showing four or six distinct objects on a floor plane (see Figure 1). For each trial, the objects were randomly chosen from a set of 16 objects. Objects were scaled to fit within a bounding box of 1.5 x 1.5 x 1.5 units within our virtual coordinate system (degrees of visual angle in memory image: width: 1.7 - 3.0; height: 1.7 - 2.7; depth: 0.5 - 1.3). The grey part of the floor plane measured 13 x 13 units (degrees of visual angle in memory image: width: 15.5 - 25.4; height: 7.0). Objects were positioned randomly within the grey part of the floor plane with a minimum center-to-center distance of 2.5 units between any two objects. In contrast to previous research (e.g., Jiang et al., 2000), we used distinct objects instead of identical objects in order to support the establishment of object correspondence across the scene rotations (Papenmeier, Meyerhoff, Jahn, & Huff, in press). A colored part (RGB: 174, 159, 84) was added to the rear site of the floor plane in order to disambiguate the direction of scene rotations. Participants viewed the scene from a viewpoint 20° above the floor plane. We used the outer buttons of a DirectIN High-Speed Button-Box to register participants' yes/no responses (button-assignment balanced across participants).

Procedure and Design

We used a location change detection paradigm with the following timing: memory image (six objects on floor plane) for two seconds, black screen for one second, probe image shown until response. The probe image depicted either six or four objects. The viewpoint on the scene either remained the same as in the memory image or the scene was rotated by 30° or 60° (left vs. right direction counterbalanced). One object in the probe image was highlighted by a red circle underneath it. Participants responded whether the location of the highlighted object on the floor plane had changed or not. In half the trials, the probed object was displaced to a random valid position on the floor plane that was not occupied by any object in the memory image. All other objects never changed locations. Analogous to previous research (e.g., Papenmeier et al., 2012), we instructed participants to memorize the locations of each object individually and to ignore the configuration formed by the objects. Participants were instructed to respond as accurately as possible and that accuracy was more important than response time.

This resulted in a 2 (spatial configuration: full, partial) x 3 (viewpoint change: 0°, 30°, 60°) x 2 (change present: yes, no) x 16 (repetitions) within-subjects design with 192 experimental trials preceded by 24 practice trials. Presentation order was randomized and practice trial conditions were balanced according to the experimental design.

Results and Discussion

The analyses are based on the signal detection theory. We report the sensitivity measure d' as dependent variable for location change detection performance. Because d' is not defined for hit rates and false alarm rates of 1.0 and 0.0, we adjusted such values to half a trial incorrect. All trials with response times larger than 8,000 ms were considered invalid and were removed from the data set (13 trials, 0.23% of the data).

We analyzed our data with linear-mixed effects models (lme; Pinheiro et al., 2013). This allowed us to treat viewpoint change as a continuous variable with our within-subjects design. In order to determine the effects of spatial configuration, viewpoint change, and the interaction of spatial configuration and viewpoint change on location change detection performance (see Figure 2A), we fitted lme models – all including a random intercept for the participant-effect and sensitivity as dependent measure – by maximum likelihood. We used likelihood-ratio tests to determine whether the stepwise inclusion of the above variables to the lme model increased the model fit significantly, thus suggesting a significant effect of the respective variables. We started with an lme model that included the intercept of the whole model as fixed effect only. The stepwise inclusion of the variables spatial configuration, $\chi^2(1) = 7.14$, $p = .008$, viewpoint change, $\chi^2(1) = 93.69$, $p < .001$, and the interaction of spatial configuration and viewpoint change, $\chi^2(1) = 5.15$, $p = .023$, each caused a significant increase of the model fit. Most importantly, the significant interaction of spatial configuration and viewpoint change suggests a viewpoint-dependent representation of contextual information in VWM. That is, the effect of context on location change detection performance decreased with increases in viewpoint change.

We ran an additional analysis to ensure that the variable viewpoint change can indeed be treated as continuous instead of discrete. Basically, treating it as continuous suggests a linear decrease of performance across increases in viewpoint change and that the slope of this decrease differs between the full spatial context and partial spatial context conditions. In contrast, treating it as discrete does not make a linearity assumption and can also describe more complex patterns. For this analysis, we fit another lme model similar to the final model above but with viewpoint change treated as discrete variable. The comparison of these two lme models treating viewpoint change as either discrete or continuous with a likelihood-ratio test revealed no significant difference, $\chi^2(2) = 3.86$, $p = .145$. That is, treating viewpoint change as a discrete variable and thereby adding additional parameters to the model did not result in a significantly increased model fit. Therefore, assuming a linear effect of viewpoint change is the more parsimonious assumption based on the present data. Future research should explore whether this linearity assumption also holds for smaller viewpoint changes.

Interestingly, with viewpoint changes of 60° there was no significant difference between the two configuration conditions, $t(29) = 1.17$, $p = .250$, while performance in both conditions was well above chance, ts >= 7.36, ps < .001. This indicates a separation of context memory and location memory in VWM. Alternatively, partial configurations might provide enough contextual information to cause the above chance performance. This was further investigated in Experiment 2 by replacing the partial spatial configurations conditions with no spatial configuration conditions that contained only the probed object but none of the other memorized objects.

Experiment 2

Method

Participants

Twenty students participated in this experiment in exchange for course credit. All participants reported normal or corrected-to-normal vision.

Apparatus and Stimuli

The apparatus was identical to Experiment 1. In contrast to Experiment 1, we removed the partial spatial configuration condition and introduced a no spatial configuration condition that contained only the probed object without the other five objects during the probe image.

Procedure and Design

We used the same procedure as in Experiment 1. The 30° viewpoint change condition was also removed from the experiment. This resulted in a 2 (spatial configuration: full, no) x 2 (viewpoint change: 0°, 60°) x 2 (change present: yes, no) x 24 (repetitions) within-subjects design with 192 experimental trials and 16 practice trials.

Results and Discussion

Change detection performance (see Figure 2B) was analyzed as in Experiment 1. One participant was removed from the analysis due to chance performance. All trials with response times larger than 8,000 ms were considered invalid and were removed from the data set (7 trials, 0.19% of the data).

As in Experiment 1, we started with an lme model that included the intercept of the whole model as fixed effect only. The stepwise inclusion of spatial configuration, $\chi^2(1) = 6.14$, $p = .013$, viewpoint change, $\chi^2(1) = 74.95$, $p < .001$, and the interaction of spatial configuration and viewpoint change, $\chi^2(1) = 16.45$, $p < .001$, each caused a significant increase of the model fit. This resembles our findings from Experiment 1 by showing a viewpoint-dependent representation of contextual information in VWM.

Importantly, with viewpoint changes of 60° there was no significant difference between the two configuration conditions, $t(18) = 0.83$, $p = .418$, and performance in both conditions was again significantly above chance, both ts ≥ 6.65, ps $\leq .001$. That is, the same pattern of results was found although the partial spatial configuration condition was replaced by no spatial configuration condition. Therefore, the alternative explanation that above chance performance was caused by the partial context available in the partial spatial configuration condition can be rejected and this experiment provides further support for a separation of context memory and location memory in VWM.

General Discussion

Previous research demonstrated the role of contextual information for the representation of individual objects in VWM (Brady & Alvarez, 2011; Hollingworth, 2007; Jiang et al., 2000, 2004; Papenmeier et al., 2012). Based on findings on contextual-cueing (Chua & Chun, 2003; Jiang et al., 2013) and a recent model of VWM (Wood, 2011), we hypothesized that the effect of contextual information, in particular spatial configurations, is viewpoint-dependent. That is, we hypothesized that location change detection performance for a single object benefits from the presence of the full spatial configuration the most without viewpoint changes. With increasing deviations in viewpoints between memory image and probe image, the effect of spatial configurations should decrease. We confirmed this hypothesis, thus providing evidence for a viewpoint-dependent representation of contextual information in VWM.

Whereas context effects occurred without viewpoint changes, context information was not utilized for change detection at large viewpoint deviations between memory and probe. Nevertheless, location change detection performance for the individual items was still well above chance at large viewpoint changes. That is, although location information for individual items can still be accessed at large viewpoint changes, contextual information cannot. This indicates that contextual information might be stored separately from the location information of individual objects, thereby supporting the notion of multiple visual storage systems within VWM (Wood, 2011). Research performed within the slot-versus-resources debate and research on effects of contextual information in working memory tasks might, therefore, focus on two different storage systems within VWM. Further research on how contextual information is stored and by which process it influences memory for individual items might, therefore, provide valuable information on the organization of VWM and thereby the relation between research on contextual information and the slot-versus-resources debate.

Note, that our findings regarding the viewpoint-dependence of contextual information in VWM account for the role of contextual information in supporting memory for individual objects. This should not be taken as evidence that participants are unable to detect spatial configurations from new viewpoints at all. Indeed, participants can detect changes of spatial layouts with virtually no errors but only increases in reaction times with increasing viewpoint changes (Diwadkar & McNamara, 1997). Shifting the focus from individual objects to a more global mode of processing (e.g., reducing memory image presentation times) might also affect the memory process and produce different results. Nonetheless, our results show that spatial configurations are not obligatorily used to facilitate the recall of individual object locations under all conditions, such as with viewpoint changes.

Our viewpoint changes were created by depth-rotations. When objects are rotated in depth the relationship between its parts remains relatively stable and thereby supports object recognition (Biederman & Bar, 1999; Biederman & Gerhardstein, 1993). Similarly with our stimuli, the spatial relations of the objects on the floor plane remained relatively stable in 3D-space across the depth-rotations, thus allowing us the study of spatial configurations. Even though this also changed object size and orientation in the 2D-projection on the screen, our objects were easily recognizable across viewpoint changes. Future research could, nonetheless, further explore the role of the visibility of local object features across the rotation (Hayward & Tarr, 1997), such as by showing canonical views of the objects (Garsoffky, Schwan, & Huff, 2009; Palmer, Rosch, & Chase, 1981) during memory and/or probe.

Future research could also explore the process by which contextual information is matched across viewpoint changes. In particular, it could be further explored how much information is matched across smaller viewpoint changes and to what extend this requires VWM capacity, for example by varying VWM load with a secondary task. Furthermore, it could be explored whether contextual information can be updated to a new viewpoint either by proprioceptive feedback when participants walk to the new viewpoint (Simons & Wang, 1998), or by a retention phase showing an empty floor plane rotating to the new viewpoint, thereby triggering spatial updating (Meyerhoff, Huff, Papenmeier, Jahn, & Schwan, 2011).

With our present experiments, we investigated contextual information in terms of the spatial configuration formed by the memorized objects. Future research should investigate viewpoint effects with other kinds of contextual information, such as the size of memorized objects (Brady & Alvarez, 2011). If contextual information is encoded as a view-dependent snapshot (Wood, 2011), similar effects for object size might be found across viewpoint changes.

Summarizing, we found evidence that contextual information supports memory for individual items in a viewpoint-dependent manner. The effect of spatial configurations on location change detection performance decreased with increases in viewpoint change between memory and probe. Futhermore, performance was well above chance at viewpoint changes of 60° where no context effects were observed. This indicates a separation of context memory and location memory in VWM.

Acknowledgements

We thank Sophia Press and Jonathan Scheeff for their help in conducting the experiment. We also thank Konstantin Sering for his help in creating the experimental stimuli and Lucy Vanes for the proofreading of this manuscript.

References

Biederman, I., & Bar, M. (1999). One-shot viewpoint invariance in matching novel objects. *Vision Research*, *39*, 2885–2899.

Biederman, I., & Gerhardstein, P. C. (1993). Recognizing depth-rotated objects: evidence and conditions for three-dimensional viewpoint invariance. *Journal of Experimental Psychology: Human Perception and Performance*, *19*, 1162–1182.

Brady, T. F., & Alvarez, G. A. (2011). Hierarchical encoding in visual working memory: Ensemble statistics bias memory for individual items. *Psychological Science*, *22*, 384–392. doi:10.1177/0956797610397956

Bülthoff, H. H., & Edelman, S. (1992). Psychophysical support for a two-dimensional view interpolation theory of object recognition. *Proceedings of the National Academy of Sciences*, *89*(1), 60–64.

Chua, K.-P., & Chun, M. M. (2003). Implicit scene learning is viewpoint dependent. *Perception & Psychophysics*, *65*, 72–80.

Chun, M. M., & Jiang, Y. (1998). Contextual cueing: Implicit learning and memory of visual context guides spatial attention. *Cognitive Psychology*, *36*, 28–71.

Diwadkar, V. A., & McNamara, T. P. (1997). Viewpoint Dependence in Scene Recognition. *Psychological Science*, *8*, 302–307. doi:10.1111/j.1467-9280.1997.tb00442.x

Garsoffky, B., Schwan, S., & Huff, M. (2009). Canonical views of dynamic scenes. *Journal of Experimental Psychology: Human Perception and Performance*, *35*, 17–27. doi:10.1037/0096-1523.35.1.17

Hayward, W. G., & Tarr, M. J. (1997). Testing conditions for viewpoint invariance in object recognition. *Journal of Experimental Psychology: Human Perception and Performance*, *23*, 1511–1521.

Hollingworth, A. (2007). Object-position binding in visual memory for natural scenes and object arrays. *Journal of Experimental Psychology: Human Perception and Performance*, *33*, 31–47. doi:10.1037/0096-1523.33.1.31

Huff, M., Papenmeier, F., Jahn, G., & Hesse, F. W. (2010). Eye movements across viewpoint changes in multiple object tracking. *Visual Cognition*, *18*, 1368–1391. doi:10.1080/13506285.2010.495878

Jiang, Y., Chun, M. M., & Olson, I. R. (2004). Perceptual grouping in change detection. *Perception & Psychophysics*, *66*, 446–453.

Jiang, Y., Olson, I. R., & Chun, M. M. (2000). Organization of visual short-term memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *26*, 683–702. doi:10.1037//0278-7393.26.3.683

Jiang, Y., Swallow, K. M., & Capistrano, C. G. (2013). Visual search and location probability learning from variable perspectives. *Journal of Vision*, *13*(6), 1–13. doi:10.1167/13.6.13

Meyerhoff, H. S., Huff, M., Papenmeier, F., Jahn, G., & Schwan, S. (2011). Continuous visual cues trigger automatic spatial target updating in dynamic scenes. *Cognition*, *121*, 73–82. doi:10.1016/j.cognition.2011.06.001

Palmer, S., Rosch, E., & Chase, P. (1981). Canonical perspective and the perception of objects. In J. Long & A. Baddeley (Eds.), *Attention and performance IX* (pp. 135–151). Hillsdale, NJ: Erlbaum.

Papenmeier, F., Huff, M., & Schwan, S. (2012). Representation of dynamic spatial configurations in visual short-term memory. *Attention, Perception, & Psychophysics*, *74*, 397–415. doi:10.3758/s13414-011-0242-3

Papenmeier, F., Meyerhoff, H. S., Jahn, G., & Huff, M. (in press). Tracking by location and features: Object correspondence across spatiotemporal discontinuities during multiple objekt tracking. *Journal of Experimental Psychology: Human Perception and Performance*.

Pinheiro, J., Bates, D., DebRoy, S., Sarkar, D., EISPACK authors, & R-core. (2013). nlme: Linear and Nonlinear Mixed Effects Models (Version 3.1-109). R package version 3.1-109.

Simons, D. J., & Wang, R. F. (1998). Perceiving real-world viewpoint changes. *Psychological Science*, *9*, 315–320.

Wood, J. N. (2011). A core knowledge architecture of visual working memory. *Journal of Experimental Psychology: Human Perception and Performance*, *37*, 357–381. doi:10.1037/a0021935

Woodman, G. F., Vogel, E. K., & Luck, S. J. (2012). Flexibility in visual working memory: Accurate change detection in the face of irrelevant variations in position. *Visual Cognition*, *20*, 1–28. doi:10.1080/13506285.2011.630694

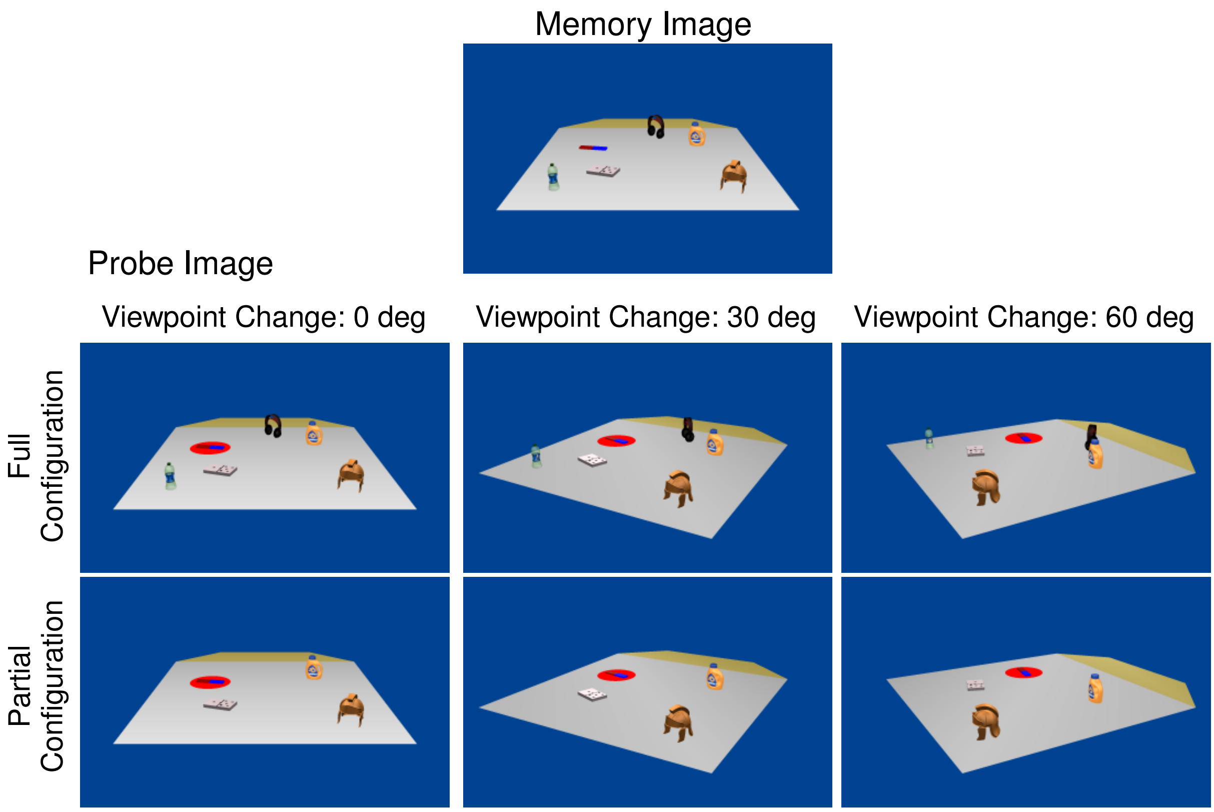

Figure 1. Stimuli used in Experiment1. Participants performed a location change detection task for one object highlighted during the probe image (figure shows target trial). Viewpoint and contextual information were manipulated.

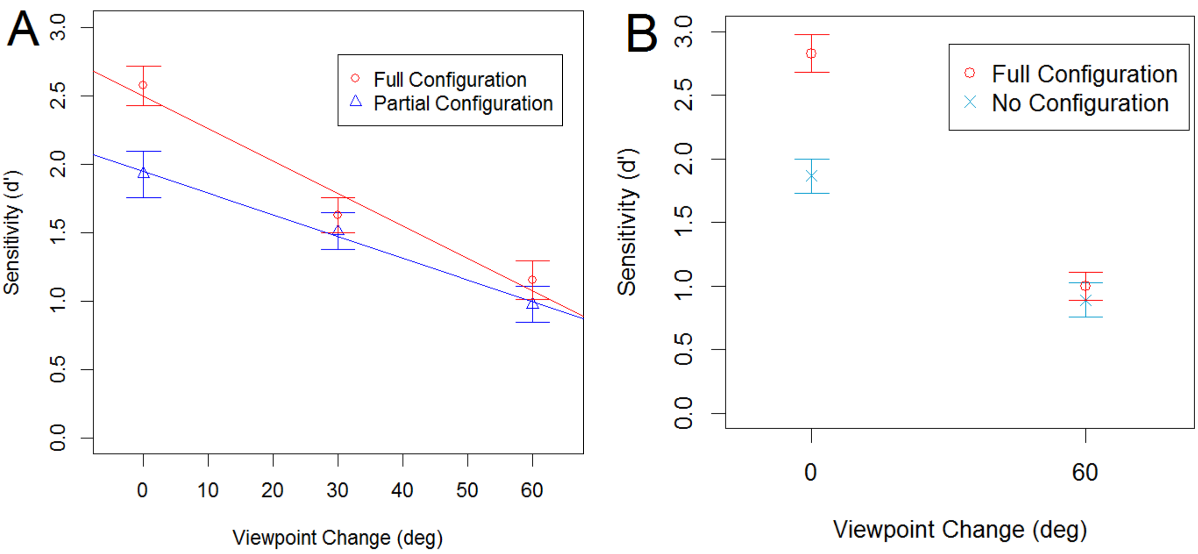

Figure 2. Results of A) Experiment 1, B) Experiment 2. The benefit of spatial configuration on performance decreased with increases in viewpoint change, suggesting a viewpoint-dependent representation of contextual information in VWM. Lines represent the prediction of the lme model (see Results section). Error bars represent the SEM.