



Breaking the cardinal rule: The impact of interitem interaction and attentional priority on the cardinal biases in orientation working memory

Gi-Yeul Bae¹

Accepted: 30 August 2021 / Published online: 17 October 2021
© The Psychonomic Society, Inc. 2021

Abstract

Although it is not typically assumed in influential models of visual working memory (WM), representations in WM are systematically biased by multiple factors. Orientation representations are biased away from the cardinal axis (i.e., cardinal bias) and they are biased away from or toward the other orientation simultaneously held in WM (i.e., interitem interaction). The present study investigated the extent to which these two bias mechanisms interact in WM. In Experiment 1, participants remembered two sequentially presented orientations and reproduced both orientations after a short delay. Cardinal biases were assessed separately for the trials where the two mechanisms produce biases in the same direction (i.e., congruent trials) and the trials where they produce biases in the opposite direction (i.e., incongruent trials). Whereas congruent trials exhibited a typical cardinal bias, incongruent trials exhibited no cardinal bias, demonstrating that the cardinal bias was canceled out by the interitem interaction. Follow-up experiments extended these results by manipulating attentional priority for the two orientations by means of precue (Experiment 2) and postcue (Experiment 3). In both experiments, attentionally prioritized items exhibited a typical cardinal bias irrespective of the congruency whereas attentionally unprioritized items exhibited a reversal of the cardinal bias in the incongruent trials, demonstrating that selective attention modulates the influence of the interitem interaction. Together, these results suggest that WM leverages information about specific stimuli and their relationship to support a given behavioral goal.

Keywords Cardinal bias · Interitem interaction · Selective attention · Visual working memory

Visual working memory (WM) provides an online storage of visual information that is relevant to the current goal of behavior. Although prevailing models of visual WM typically assume that different stimulus values are stored in WM with equal representational profile (van den Berg et al., 2012; Zhang & Luck, 2008), increasing number of studies demonstrated that visual WM behavior systematically varies in a stimulus-specific manner. For example, color WM is biased away from the nearest category boarder (Bae et al., 2015), and orientation WM is biased away from the nearest cardinal axis (Bae & Luck, 2019; Pratte et al., 2017; Wei & Stocker, 2015). When multiple items are simultaneously held in WM, the interitem interaction biases individual WM representations depending on their similarity. For example, when two colors are held in WM, a report of one color is biased away from the

other color if the two colors are similar and the report is biased toward the other color if they are dissimilar (Golomb, 2015). The same pattern of the results was observed when two orientations were held in WM even when perceptual interaction between them was prevented (Bae & Luck, 2017).

Although the two mechanisms for WM biases—categorical structure of the stimulus space and the interitem interaction—operate simultaneously in many WM tasks, they were studied independently and it is not well understood how the two mechanisms would impact individual WM representations. To answer this question, the present study focused on the case in which the expected direction of biases from the two mechanisms were in conflict. For example, when two orientations are stored in WM, one orientation will be biased away from the nearest cardinal axis (i.e., the cardinal bias), but, at the same, the same orientation will be biased toward the nearest cardinal axis if the other orientation repels the orientation away from it (i.e., interitem interaction). What would happen to a WM representation when it is repelled by the two bias mechanisms in the opposite direction?

✉ Gi-Yeul Bae
giyeulbae@asu.edu

¹ Department of Psychology, Arizona State University, 950 S. McAllister Ave, Tempe, AZ 85287, USA

One possibility is that WM representations are categorically biased irrespective of the interitem interaction. If our visual system follows ‘cardinal rule’ as a general principle when representing orientations (Girshick et al., 2011), then the categorical structure of the stimulus space (i.e., cardinal axis) may play a dominant role by serving an absolute reference frame for the orientations. Another possibility is that WM representations are biased by the interitem interaction irrespective of the cardinal axis of the orientation space. It is well known that the visual system incorporates a mechanism that separate one representation from the other when processing multiple items (e.g., lateral inhibition; Blakemore et al., 1970) and studies have shown that the mechanism indeed underlies perceptual interaction between orientations (“tilt-illusion”; Gibson & Radner, 1937). Given the recent finding that the similar mechanism operates on WM representations as well (Kiyonaga & Egner, 2016), it is possible that a given orientation representation would be predominantly influenced by the other orientation held in WM. Alternatively, it is also possible that the two mechanisms simultaneously impact orientation representations. If this is the case, then the cardinal bias would be reduced or cancelled out when the interitem interaction biases a given orientation representation in the direction opposite to the cardinal bias.

To test these possibilities, I had participants perform an orientation WM task in which they remembered orientations of two sequentially presented teardrop objects and reported each orientation in a random order (see Fig. 1a). Previous study has demonstrated that the two orientations in WM interact each other, exhibiting systematic biases depending on the similarity between them (Bae & Luck, 2017). Specifically, Bae and Luck (2017) found that the reported orientation was biased away from the other orientation in WM when the two orientations were similar (e.g., $\Delta < 90^\circ$) whereas the report was biased toward the other item when the two orientations were dissimilar (e.g., $\Delta > 90^\circ$). The present study investigated how this interitem interaction influences the cardinal bias for each item. In the analysis, I assessed the cardinal bias separately for the trial in which the interitem interaction and the nearest cardinal axis induce biases in the same direction (i.e., congruent trials) and for the trials in which they induce biases in the opposite direction (i.e., incongruent trials). Examples of these trials are illustrated in Fig. 1b.

To foreshadow the main findings, the cardinal bias was completely eliminated when the interitem interaction induced biases in the opposite direction of the cardinal bias. In two follow-up experiments, I investigated how attentional priority might play a role in the interplay between the cardinal bias and the interitem interaction. I found that attentionally prioritized item still exhibited cardinal biases even in the face of the impact of the interitem interaction and attentionally unprioritized item exhibited the reversal of the cardinal bias. These results demonstrate that the two bias mechanisms

contribute to the representational biases in WM in an attention-dependent manner, providing an accountable source of trial-to-trial variability in WM behavior.

Experiment 1

Method

This article reports new analyses of previously published data (Bae & Luck, 2017). All the critical information about the stimuli, task, and analyses is reported here. Additional details can be found in that article.

Participants Sixteen college students (nine females; age range: 18–30 years) with normal or corrected-to-normal visual acuity participated for course credit after providing informed consent. The sample size was determined a priori on the basis of previous research that utilized delayed estimation procedure. The study was approved by the Institutional Review Board of Arizona State University.

Stimuli and procedure Stimuli were presented on a Dell U2412M LCD monitor with a gray background (31.2 cd/m^2) at a viewing distance of 70 cm. A black fixation dot was continuously present except during the intertrial interval.

Figure 1a depicts the task procedure. Each trial began with the fixation dot. After 500 ms, the first target was presented for 200 ms, followed by a 750-ms blank interval. The second target was then presented for 200 ms. Participants were asked to remember the orientations of the two teardrop objects as precisely as possible. The orientations on a given trial were randomly chosen from 16 equally spaced values (separated by 22.5° , starting at 11.25° from upright) with the constraint that they were never identical. Thus, the orientation difference between the two targets could be $\pm 22.5^\circ$, $\pm 45^\circ$, $\pm 67.5^\circ$, $\pm 90^\circ$, $\pm 112.5^\circ$, $\pm 135^\circ$, $\pm 157.5^\circ$, or 180° . After a 1,000 ms blank interval, a response ring appeared along with the text “first orientation” or “second orientation” at the top of the screen, indicating which target the observer should report. Participants reproduced the specified target orientation using a computer mouse. After the first report, there was a 500-ms gap, and then a second response ring appeared along with an instruction to report the other target. The order of report was randomized. Participants performed the next trial after 1-s delay. Each participant completed four blocks of 64 trials after completing 16 practice trials.

Analysis For each trial, response error was computed by taking the angular difference between actual target orientation and the reported orientation. Critically, positive sign was given to response errors away from the nearest cardinal orientation and negative sign was given to response errors toward the

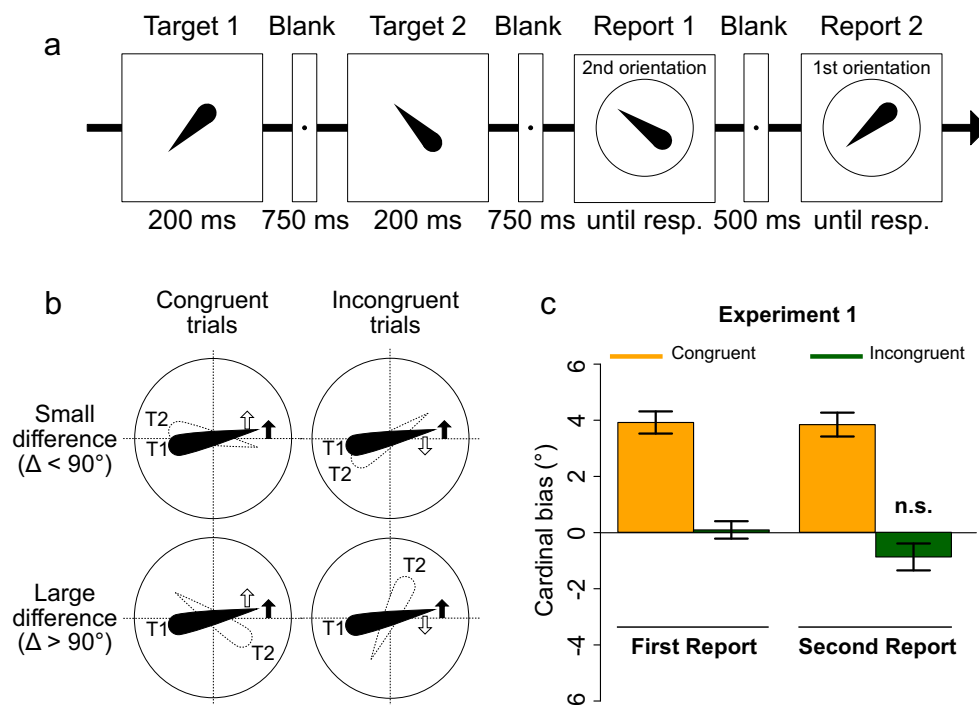


Fig. 1 **a** Orientation delayed estimation. Participants remembered the orientations of two sequentially presented teardrop items and reported each orientation in an order indicated by a cue (either first orientation or second orientation). The order of report was randomized across trials. **b** Congruency between the direction of the bias driven by the categorical structure of the orientation space (i.e., black arrow) and the bias driven by the other orientation in WM (i.e., white arrow). In this example, the orientation of black teardrop should be reported and the orientation of

white teardrop represents the other orientation held in WM. In congruent trials, the cardinal axis (i.e., horizontal line) and the other orientation bias the target orientation in the same direction. In incongruent trials, however, the cardinal axis and the other orientation bias the same target orientation in the opposite direction. **c** Cardinal biases for congruent and incongruent trials separated by the order of report in Experiment 1. Positive value indicates that the reports were biased away from the nearest cardinal axis (i.e., a typical cardinal bias). Error bars indicate $\pm 1 SE$

nearest cardinal orientation. Thus, positive error indicates typical cardinal biases observed in previous studies (Bae & Luck, 2019; Pratte et al., 2017; Wei & Stocker, 2015).

To examine how the cardinal biases would be influenced by the interitem interaction, data was grouped in terms of the congruency between the cardinal bias and the bias driven by the interitem interaction (see Fig. 1b). When the orientation difference is less than 90° (i.e., $\pm 22.5^\circ$, $\pm 45^\circ$, $\pm 67.5^\circ$), a given target orientation near the cardinal orientation (e.g., 78.75°) would be repelled by the nearest cardinal orientation (i.e., 90°), and that orientation would also be repelled by the other orientation if the other orientation is in the same side as the cardinal orientation (e.g., 101.25° ; see Fig. 1b, top-left). This type of trials was grouped as *congruent* trials. However, if the near-cardinal orientation (e.g., 78.75°) was paired with the other orientation with the same orientation difference but in the opposite side of the cardinal axis (e.g., 56.25°), then the other orientation would repel the target orientation toward the cardinal orientation (see Fig. 1b, top-right). This type of trials was grouped as *incongruent* trials (see Fig. 1b).

When the orientation difference is larger than 90° (i.e., $\pm 112.5^\circ$, $\pm 135^\circ$, $\pm 157.5^\circ$), a given target orientation near the cardinal orientation (e.g., 78.75°) would be repelled by the nearest cardinal orientation (i.e., 90°) and that orientation

would be attracted toward the other orientation (e.g., 326.25°) if the other orientation is in the same side as the target from the cardinal orientation (*congruent* trials, Fig. 1b, bottom-left). However, if the near-cardinal orientation (e.g., 78.75°) was paired with the other orientation with the same orientation difference but in the same side of the cardinal axis (e.g., 191.25°), then the other orientation would attract the target orientation toward it (*incongruent* trials, Fig. 1b, bottom-right).

All the trials was grouped into either congruent and incongruent trials. Trials with orientation differences of $\pm 90^\circ$ and 180° were excluded from the analysis because the direction of the bias driven by the interitem interaction is undefined for those trials. However, the cardinal biases from those trials were served as a control condition to compare with the cardinal biases with the interitem interaction (see Results and Discussion). Trials with response error larger than 22.5° (i.e., smallest orientation differences tested in the experiment) were excluded from the analysis because those trials might reflect swap errors¹(Bays et al., 2009)and/or random guesses

¹ Although excluding errors larger than the smallest orientation difference does not completely remove swap errors, they can only produce attraction biases, which should decrease the overall repulsion bias observed in the present study.

(Zhang & Luck, 2008). If only one of the reports was in the exclusion criteria on a given trial, then only that report was excluded from the analysis. This exclusion criterion removed 12.9% of the total reports.

The main analysis focused on the mean of the signed response errors to assess cardinal biases for the congruent and incongruent trials. Although assessing the effect of response order was not the main concern for the present study, I assessed the cardinal bias for each report separately to examine whether the congruency effect was dominated by either the first or the second report. In addition, assessing the congruency effect separately for each report provides a control condition in the analyses for the following experiments (see Experiments 2 and 3).

I also conducted more sophisticated mixture model analysis (Zhang & Luck, 2008) with a bias parameter. The model can be expressed as this equation:

$$p(\hat{\theta}) = P_T * \Phi(\hat{\theta}, \theta_T + \mu, \kappa) + (1 - P_T) * \frac{1}{2\pi}. \quad (1)$$

In this model, $p(\hat{\theta})$ represents the estimated probability of orientation reports. There are three free parameters to estimate in the model. P_T represents the proportion of target-based reports. Thus, $1 - P_T$ represents the proportion of random reports (i.e., uniform distribution); κ is the concentration parameter of a von Mises distribution (Φ), which represents the precision of the reports. μ is the central tendency of the von Mises distribution (Φ) for target-based reports, representing the bias in the target-based reports. A positive μ reflects attraction toward the nearest cardinal orientation and a negative μ reflects repulsion away from the nearest cardinal orientation. Therefore, μ parameter is analogous to the signed response error in the main analyses. Maximum likelihood estimates of each parameter were obtained using a nonlinear optimization algorithm (Nelder & Mead, 1965), separately for each participant. To ensure that the estimates reflect global maxima, the optimization was repeated with multiple different initial parameter values. Because this model captures random guesses using a separate parameter, the model fit was done without excluding data. Overall, the mixture model analysis produced the same pattern of results as the mean of the signed response error (see [Supplementary Material](#)).

Results and discussion

The analysis for the interaction between the two orientations in this experiment is reported in Bae and Luck (2017). The present study reports how the interitem interaction impacts the cardinal biases. Figure 1c shows the mean of the signed response errors as a function of the congruency, separated by the response order. Positive errors indicate the reports were biased away from the nearest cardinal axis. When the direction of

bias from the other orientation was consistent with the direction of the bias from the nearest cardinal axis (i.e., congruent trials), typical cardinal bias was observed irrespective of the report order. However, when the direction of bias from the other orientation was not consistent with the direction of the bias from the nearest cardinal axis (i.e., incongruent trials), the cardinal bias was no longer present. Instead, reports were slightly biased toward the nearest cardinal axis in the trials. This was statistically tested using a two-way analysis of variance (ANOVA), with factors of congruency (congruent or incongruent) and response order (Report 1 and Report 2). The main effect of congruency was significant, $F(1, 60) = 110.11, p < .001, \eta_p^2 = .65$. However, neither the main effect of response order, $F(1, 60) = 1.62, p = .208, \eta_p^2 = .03$, nor the two-way interaction, $F(1, 60) = 1.19, p < .280, \eta_p^2 = .020$, was significant.

I conducted two sets of follow-up analyses, each using the false discovery rate (FDR) correction for multiple comparisons with an alpha level of 0.05 (Benjamini & Hochberg, 1995). The first analyses used paired t tests to compare the congruency effects for the first and the second report. The difference between congruent and incongruent trials was significant for both first, $t(15) = 10.11, p < .001, d = 2.53$, and second report, $t(15) = 6.25, p < .001, d = 1.56$, indicating that the interitem interaction influenced the cardinal biases for both reports. The second analyses used one-sample t test to compare the cardinal bias against zero for each combination of the congruency and the order of report. Congruent trials produced significant cardinal bias for both first, $t(15) = 9.93, p < .001, d = 2.48$, and second, $t(15) = 9.02, p < .001, d = 2.25$, reports. However, incongruent-first reports exhibited no significant cardinal bias, $t(15) = .27, p = .788, d = .07$, indicating that the cardinal bias was cancelled out by the interaction between the two orientations. Incongruent-second reports exhibited some hint of the reversal of the cardinal bias but it was not statistically significant, $t(15) = 1.83, p = .088, d = .46$. Together, these results are consistent with the hypothesis that both the cardinal axis and the interitem interaction contribute to the representational biases for the orientations in WM.

To investigate whether the cardinal bias in the congruent trials was a result of the additive effect of the cardinal axis and the interitem interaction, I compared the cardinal biases between the congruent trials and the trials with no meaningful interitem interaction (i.e., orientation differences of $\pm 90^\circ$ and 180°). For both first and second reports, the congruent trials produced numerically greater cardinal biases compared with the trials without interitem biases (Congruent/First report: 3.92° vs. No-interitembiases/First report: 3.25° ; Congruent/Second report: 3.84° vs. No-interitembiases/Second report: 3.45°). However, the differences were not statistically significant, First reports: $t(15) = 1.62, p = .13, d = 0.41, BF_{01} = 1.34$; Second reports: $t(15) = .93, p = .37, d = .23, BF_{01} = 2.70$ (see [Supplementary Material](#) for the consistent results from the

mixture model without the trial exclusion). This result may indicate that the cardinal axis and interitem interaction asymmetrically contributed to the representational biases in a manner specific to the stimulus values and their configurations. Alternatively, it is also possible that they contributed to the biases in an additive manner, but the additive effect was not significant due to the lack of statistical power. I discuss this issue in more detail in General Discussion.

Experiment 2

Experiment 1 demonstrated that the cardinal bias was completely eliminated in the incongruent-first reports. Intriguingly, however, there were some hint of reversal of the cardinal bias for the incongruent-second reports. This may reflect the possibility that the item for the second report received less attentional priority so that it was less protected from the impact of the item reported first (Huang & Sekuler, 2010a; Matsukura et al., 2007). Experiment 2 directly tested this possibility by attentionally prioritizing one of the two orientations via precue. On each trial, a precue (“1” or “2”) was presented to indicate which of the two upcoming orientations should be given higher priority for that trial (see Fig. 2a). Two results were predicted. It was predicted that the uncued item should exhibit the reversal of the cardinal bias for the incongruent trials if the uncued item is less protected from the impact of the cued item (thus receives greater influence from the cued item). It was also predicted that the cued item should exhibit the cardinal bias even for incongruent trials if attention protects the cued item from the influence of the uncued item (thus receives greater influence from the cardinal axis).

Method

The methods were identical to those of Experiment 1, except as noted here. A new group of 16 college students (six females; age range: 18–30 years) participated for course credit after providing informed consent.

The procedure is illustrated in Fig. 2a. Each trial started with a digit (either “1” or “2” with equal probability) presented at the center of the screen for 200 ms, indicating which item should be prioritized in that trial. After a 750-ms blank, the two orientations were sequentially presented as in Experiment 1. After another 1,000-ms blank, the observer was serially cued to reproduce each of the two orientations.

To motivate participants to give greater attentional priority to the cued target, the task incorporated two additional manipulations. First, the observer received 5 points for an accurate response (error $\leq 10^\circ$) to the cued target but only 1 point for an accurate response to the uncued target. Second, the cued target was always reported first. The exclusion criteria used in

Experiment 1 removed 13.56% of the reports in this experiment. As in Experiment 1, I conducted a mixture model analysis (Zhang & Luck, 2008) and found that the bias parameter from the model exhibited the same pattern of results (see [Supplementary Material](#)).

Results and discussion

To confirm the effect of precue on the quality of orientation representations, I conducted a mixture model analysis and compared the estimated precision for the cued item with the precision for the uncued item. The cued item ($\kappa = 47.08$) exhibited greater precision than the uncued item ($\kappa = 24.76$), $t(15) = 4.12, p < .001, d = 1.03$. To ensure that this effect was not merely driven by the differences in the response order, I compared the precision difference between cued and uncued reports in this experiment with the precision difference between first and second reports in Experiment 1 using an independent-samples *t* test. I found that the precue effect was significantly greater than the response order effect, $t(30) = 2.98, p = .005, d = 1.05$, demonstrating that the cued item in this experiment received greater attentional priority.

Figure 2c shows the cardinal bias as a function of the congruency, separated by precue. For the congruent trials, a typical cardinal bias was observed irrespective of precue. For the incongruent trials, the cued reports exhibited the cardinal bias whereas the uncued reports exhibited the reversal of the cardinal bias. This was statistically tested using a two-way ANOVA, with factors of congruency (congruent or incongruent) and precue (cued and uncued). The main effect of congruency was significant, $F(1,60) = 41.11, p < .001, \eta_p^2 = .41$. The main effect of precue was significant, $F(1, 60) = 6.75, p = .012, \eta_p^2 = .10$. Importantly, the two-way interaction was significant, $F(1, 60) = 11.96, p = .001, \eta_p^2 = .17$.

To further investigate these results, I conducted follow-up analyses using FDR correction for multiple comparisons (Benjamini & Hochberg, 1995). Both congruent-cued reports, $t(15) = 7.06, p < .001, d = 1.76$, and incongruent-cued reports, $t(15) = 5.732, p < .001, d = 1.43$, exhibited significant cardinal biases. However, the cardinal bias was significantly greater for the congruent-cued reports than for the incongruent-cued reports, $t(15) = 3.32, p = .005, d = .83$. These results indicate that the attentional prioritization protected the cued item but did not completely eliminate the impact of the uncued item. The congruency effect was significant for the uncued reports, $t(15) = 7.15, p < .001, d = 1.79$. Congruent-uncued reports exhibited a significant cardinal bias, $t(15) = 4.99, p < .001, d = 1.25$. However, incongruent-uncued reports exhibited a significant reversal of the cardinal bias, $t(15) = -3.19, p = .006, d = .80$. These results are consistent with the prediction that attentionally prioritized item was more protected from the impact of the unprioritized item (thus received more impact from the cardinal axis) but unprioritized item was less

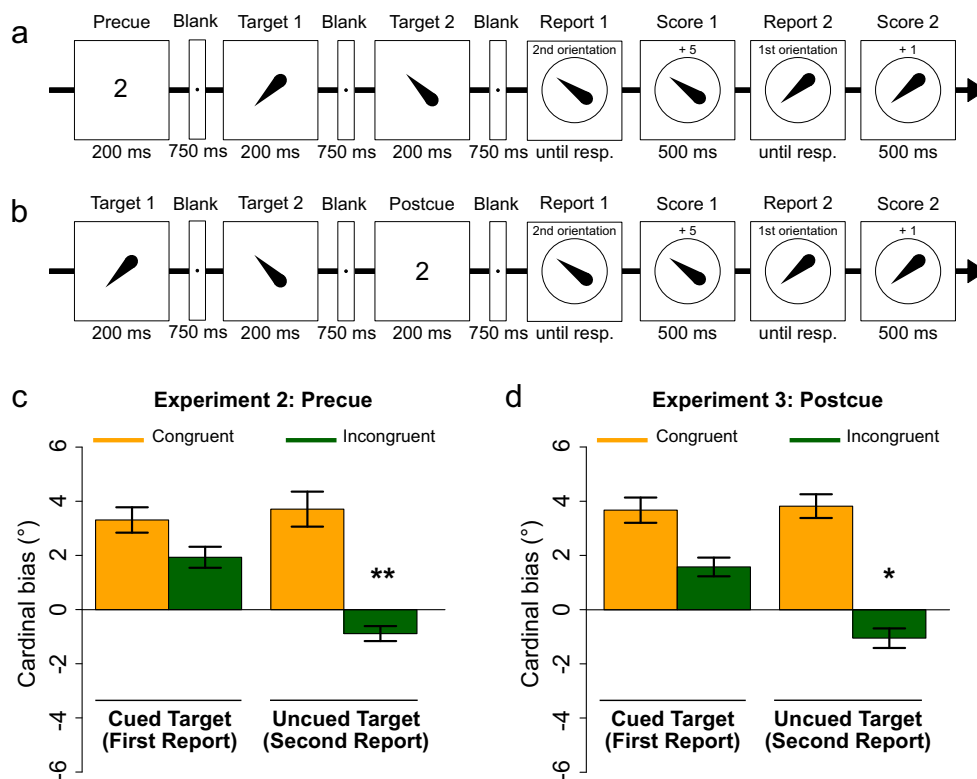


Fig. 2 **a** Orientation delayed estimation with precue used in Experiment 2. Each trial began with a precue indicating which of the two upcoming orientations should be prioritized. The two targets were then presented, followed by report of the two orientations. To motivate the attentional prioritization, the cued orientation was always reported first and participants received 5 points as a feedback for an accurate report (error $\leq 10^\circ$) of the cued item but 1 point for an accurate report of the uncued

item. In this example, the second target is cued and reported first. **b** Orientation delayed estimation with postcue used in Experiment 3. A postcue was presented after the presentation of the second target. **c–d** Cardinal biases for congruent and incongruent trials separated by the cueing condition in Experiment 2 and Experiment 3. Positive value indicates the reports were biased away from the nearest cardinal axis. Error bars indicate ± 1 SE. ** $p < .01$, * $p < .05$

protected from the impact of the prioritized item (thus received more impact from the cued item).

Note that Experiment 1 demonstrated that the order of report has little or no impact on the performance (see Fig. 1c). Thus, it is unlikely that the significant cardinal bias for the cued-incongruent trials was driven by the mere fact that it was always reported first. To further demonstrate this, I compared the bias for the incongruent-cued reports in this experiment with the bias for the incongruent-first-reports in Experiment 1 using an independent-samples *t* test. The difference between them was significant, $t(30) = 3.72, p < .001, d = 1.32$.

As in Experiment 1, I examined whether the cardinal bias in the congruent trials was a result of the additive effect of the cardinal axis and the interitem interaction by comparing the cardinal biases between the congruent trials and the bias for the trials with no meaningful interitem interaction (i.e., orientation differences of $\pm 90^\circ$ and 180°). In contrast to Experiment 1, the congruent trials produced significantly greater cardinal biases compared to the trials without interitem biases: (Congruent/Cued: 3.31° vs. No-interitem bias/Cued: 2.49°), $t(15) = 2.15, p = .048, d = .54$, (Congruent/Uncued: 3.71°

vs. No-interitem bias/Uncued: 2.24°), $t(15) = 2.66, p = .018, d = .67$. These results suggest that the two bias mechanisms additively impacted the representational biases in congruent trials.

Experiment 3

Experiment 2 used a precue to attentionally prioritize one item over the other. Thus, the prioritization process was likely to operate during the encoding of the stimulus. To investigate whether the attentional priority given to the representations already held in WM would exhibit similar effects, Experiment 3 used a postcue after the presentation of the two orientations (see Fig. 2b).

Method

The methods were identical to those of Experiment 2, except as noted here. A new group of sixteen college students (nine females; age range: 18–30 years) participated for course credit after giving informed consent. The procedure is illustrated in

Fig. 2b. After the encoding of the two sequentially presented stimulus, a digit (“1” or “2” with equal probability) was presented at the center of the screen for 200 ms, serving as a postcue that indicated which of the two previously presented orientations should be prioritized on that trial. The effect of postcue has been confirmed in the previous study by showing greater representational precision for the cued item versus uncued item (Bae & Luck, 2017). The exclusion criteria used in Experiment 1 removed 15.95% of the reports in this experiment. As in Experiment 1, I conducted a mixture model analysis (Zhang & Luck, 2008) and found that the bias parameter from the model exhibited the same pattern of results (see [Supplementary Material](#)).

Results and discussion

A mixture model analysis confirmed that the cued item ($\kappa = 35.34$) exhibited greater precision compared to the uncued item ($\kappa = 22.88$), $t(15) = 4.50$, $p < .001$, $d = 1.13$. This effect was significantly greater than the precision difference between the first and the second reports in Experiment 1, $t(30) = 2.20$, $p = .036$, $d = .78$, suggesting that the greater precision for the cued item was not merely driven by the differences in the report order.

Figure 2d shows the cardinal bias as a function of the congruency, separated by the postcue. The pattern of results was consistent with the main results in Experiment 2. For the congruent trials, a typical cardinal bias was observed irrespective of postcue. For the incongruent trials, the cued reports exhibited the cardinal bias whereas the uncued reports exhibited the reversal of the cardinal bias. This was statistically tested using a two-way ANOVA with factors of congruency (congruent or incongruent) and postcue (cued and uncued). The main effect of congruency was significant, $F(1, 60) = 73.82$, $p < .001$, $\eta_p^2 = .55$. The main effect of precue was significant, $F(1, 60) = 9.34$, $p = .003$, $\eta_p^2 = .14$. The two-way interaction was also significant, $F(1, 60) = 11.71$, $p = .001$, $\eta_p^2 = .16$.

I conducted follow-up analyses using FDR correction for multiple comparisons (Benjamini & Hochberg, 1995). As in Experiment 2, the cardinal bias was significant for both congruent-cued reports, $t(15) = 7.89$, $p < .001$, $d = 1.97$, and incongruent-cued reports, $t(15) = 8.76$, $p < .001$, $d = 2.19$, but the difference between them was significant, $t(15) = 4.05$, $p = .001$, $d = 1.01$. These results indicate that the postcue protected the cued item but did not completely eliminate the impact of the uncued item. The main effect of congruency was significant for uncued reports, $t(15) = 6.99$, $p < .001$, $d = 1.75$. Congruent-uncued reports exhibited cardinal biases, $t(15) = 4.57$, $p < .001$, $d = 1.14$, but the bias was reversed for incongruent-uncued reports, $t(15) = -2.90$, $p = .011$, $d = .72$. These results replicate the main findings in Experiment 2 and further demonstrate that attention can protect an item that is already encoded in WM.

To demonstrate that the cardinal bias for the incongruent-cued reports was not driven by the mere fact that the cued item was always reported first, I compared the bias for the incongruent-cued reports with the bias for the incongruent-first-reports in Experiment 1 using an independent-samples *t* test. The difference between them was significant, $t(30) = 3.22$, $p = .003$, $d = 1.14$.

To examine whether the cardinal bias in the congruent trials was a result of the additive effect of the cardinal axis and the interitem interaction, I compared the cardinal biases for the congruent trials with the bias for the trials with the baseline trials (i.e., orientation differences of $\pm 90^\circ$ and 180°). I found no significant difference between them for both the cued reports, (Congruent: 3.66° vs. No-interitem bias: 3.79°), $t(15) = .40$, $p = .693$, $d = .10$, and uncued reports (Congruent: 3.82° vs. No-interitem bias: 3.76°), $t(15) = .09$, $p = .929$, $d = .02$.

General discussion

The present study investigated how the two sources of WM biases—the cardinal axis and the interitem interaction—influence on the orientation representations in working memory. Using a delayed estimation with two sequentially presented orientations, Experiment 1 showed that the cardinal bias for one orientation was completely eliminated when the other orientation biased the orientation in the opposite direction of the cardinal bias. Experiment 2 extended the finding by showing that the effect was modulated by attentional priority given to an item via precue. The cardinal bias was recovered when an item received greater attentional priority and was reversed when an item was unprioritized. This attentional modulation was replicated in Experiment 3 in which the attentional priority was given via a postcue to an item that is already encoded in WM. These results indicate that the extent to which the two mechanisms interact is dependent on the selective attention. Consistent with the hypothesis that attention protects a representation from potential interferences (e.g., Matsukura et al., 2007), attentionally prioritized item was less influenced by the interitem interaction, but attentionally unprioritized item was more influenced by the interitem interaction. Although the present study does not provide evidence that the cardinal axis and the interitem interaction cooperate in a single trial level, the overall results provide converging evidence that WM utilizes information about specific stimuli and their configurations to support the current behavioral goal, rather than maintains individual items independently.

While the present study clearly showed that the cardinal bias was cancelled out by the interitem interaction in incongruent trials, it was less clear whether the cardinal bias was exaggerated by the additive effect of the interitem interaction in congruent trials. Experiments 1 and 3 found no significant

difference in the magnitude of the cardinal bias between the congruent and the no-interaction baseline trials. However, Experiment 2 found a greater cardinal bias for the congruent trials compared with the baseline trials. These inconsistent results might be due to the lack of statistical power for a given experiment. In a follow-up analysis, I examined the additive effect by collapsing the data across the three experiments and the two reports. Although the effect size was small, results showed that the cardinal bias for the congruent trials was significantly greater than the bias for the baseline trials (Congruent: 3.71° vs. No-interitem bias: 3.17°), $t(47) = 2.54$, $p = .015$, $d = .37$. Together, these results suggest that the interitem interactions amplified the cardinal bias in the congruent trials, but the effect was weaker than in the incongruent trials such that it required more statistical power. Although collapsing data across different experiments might not be an ideal way to test the additive effect of cardinal axis and interitem interaction for congruent trials, the significant effect is at least consistent with the hypothesis that the interitem interaction exaggerated the cardinal bias for the congruent trials.

It is important to ask at which processing stage the two sources of bias altered the representations. One possibility would be that the representations only exhibit the cardinal bias in WM, but response-related processes may have altered the report depending on the similarity between the representations. However, the cancellation occurred even for the first report (i.e., when no influence from the second report was present) and the order of report was randomized so that participants were not able to prepare for a report before the cue in the task. Thus, the main results cannot be attributed to the influence of the report or to the response preparation processes. Instead, it is more likely that the orientations were encoded in WM with the cardinal bias, but the interitem interaction cancelled out the cardinal bias during WM maintenance. To provide more convincing evidence for this, future study should use a neuroimaging method with high temporal resolution (e.g., EEG). Previous studies have demonstrated that visual information in WM can be decoded on the basis of the spatial pattern of EEG signals (Bae & Luck, 2018; Foster et al., 2016), and more recent work has shown that the decoded stimulus information exhibited categorical biases (Bae, 2021; Ester et al., 2020). If the decoded stimulus information from this task exhibits the cancellation of the cardinal bias even before a response is made, that would strongly support the hypothesis that the two sources of bias altered the representations in WM themselves as opposed to the reports.

The observation that WM representations are influenced by multiple sources of biases in a systematic manner has clear implication to the current theories in WM. Specifically, an influential WM model assumes that the quality of WM representations varies on a trial-by-trial basis due to the stimulus-independent factors such as stochastic neural processes and

random attentional fluctuations (e.g., van den Berg et al., 2012). However, the present study demonstrated that the magnitude of the cardinal bias was modulated by a specific combination of orientations stored in WM. This naturally suggests that WM should vary on a trial-by-trial basis if different stimulus values should be remembered in different trials (which is a common practice in most of WM studies). In other words, although WM may appear to vary across trials in a random fashion, it is likely that the significant proportion of the variability might be actually driven by specific stimulus values and the interaction between them. Indeed, a recent study found that, when the stimulus-specific variability was incorporated into WM models, the model without the random trial-to-trial variability assumption outperformed the model with the trial-to-trial variability assumption which otherwise produced superior performance (Pratte et al., 2017). Therefore, WM models should take the stimulus-specific variabilities into account otherwise the stimulus-specific variabilities can lead to unwarranted advantage for the models that assume random trial-to-trial variability.

In contrast to previous studies that showed attractive biases between similar stimulus values (e.g., Brady & Alvarez, 2011; Huang & Sekuler, 2010b), the present data showed repulsive biases for small orientation differences. Although the present study was not designed to investigate this discrepancy, it suggests that this discrepancy maybe driven by specific task demands. For example, when WM needs to maintain multiple items in a more stable manner, then it may utilize “summary” statistics when encoding individual items, thereby producing attraction biases (Brady & Alvarez, 2011). However, when WM needs to individuate individual items because each item has to be reported (as in the present study), it may exaggerate the dissimilarity between items thereby producing repulsion biases. It would be an important future research to investigate this possibility and how different task demands would distort WM representations more broadly (see Fukuda et al., [in press](#)).

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.3758/s13414-021-02374-2>.

Authors' contributions G.B. designed the study, collected and analyzed data, and wrote the manuscript.

Funding No funding was received to assist with the preparation of this manuscript.

Data availability The data and materials will be made available at <https://osf.io/vygr3> upon acceptance of the manuscript.

Declarations

Conflicts of interest/Competing interests The author has no conflicts of interest to declare that are relevant to the content of this article.

Ethics approval Approval was obtained from the ethics committee of Arizona State University. The procedures used in this study adhere to the tenets of the Declaration of Helsinki.

Consent to participate Informed consent was obtained from all individual participants included in the study.

Consent to publish All the participants has consented to the submission of their data to the journal.

Code availability Freely available statistical software (R) was used for the data analyses.

References

- Bae, G.-Y. (2021). Neural evidence for categorical biases in location and orientation representations in a working memory task. *NeuroImage*, 240(15), 118–366. <https://doi.org/10.1016/j.neuroimage.2021.118366>
- 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>
- Bae, G.-Y., & Luck, S. J. (2018). Dissociable decoding of spatial attention and working memory from EEG oscillations and sustained potentials. *The Journal of Neuroscience*, 38(2), 409–422. <https://doi.org/10.1523/JNEUROSCI.2860-17.2017>
- Bae, G.-Y., & Luck, S. J. (2019). What happens to an individual visual working memory representation when it is interrupted? *British Journal of Psychology*, 110(2), 268–287. <https://doi.org/10.1111/bjop.12339>
- Bae, G.-Y., Olkkonen, M., Allred, S. R., & Flombaum, J. I. (2015). Why some colors appear more memorable than others: A model combining categories and particulars in color working memory. *Journal of Experimental Psychology: General*, 144(4), 744–763. <https://doi.org/10.1037/xge0000076>
- 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), 7–7. <https://doi.org/10.1167/9.10.7>
- Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society: Series B (Methodological)*, 57(1), 289–300. <https://doi.org/10.1111/j.2517-6161.1995.tb02031.x>
- Blakemore, C., Carpenter, R. H., & Georgeson, M. A. (1970). Lateral inhibition between orientation detectors in the human visual system. *Nature*, 228(5266), 37–39. <https://doi.org/10.1038/228037a0>
- Brady, T., & Alvarez, G. (2011). Hierarchical encoding in visual working memory: Ensemble statistics bias memory for individual items. *Psychological Science*, 22(3), 384–392.
- Ester, E. F., Sprague, T. C., & Serences, J. T. (2020). Categorical biases in human occipitoparietal cortex. *Journal of Neuroscience*, 40(4), 917–931. <https://doi.org/10.1523/JNEUROSCI.2700-19.2019>
- Foster, J. J., Sutterer, D. W., Serences, J. T., Vogel, E. K., & Awh, E. (2016). The topography of alpha-band activity tracks the content of spatial working memory. *Journal of Neurophysiology*, 115(1), 168–177. <https://doi.org/10.1152/jn.00860.2015>
- Fukuda, K., Pereira, A. E., Saito, J. M., Tang, T. Y., Tsubomi, H., & Bae, G. Y. (in press). Working memory content is distorted by its use in perceptual comparisons. *Psychological science*. <https://doi.org/10.1177/09567976211055375>
- Gibson, J. J., & Radner, M. (1937). Adaptation, after-effect and contrast in the perception of tilted lines. I: Quantitative studies. *Journal of Experimental Psychology*, 20(5), 453–467. <https://doi.org/10.1037/h0059826>
- Girshick, A. R., Landy, M. S., & Simoncelli, E. P. (2011). Cardinal rules: Visual orientation perception reflects knowledge of environmental statistics. *Nature Neuroscience*, 14(7), 926–932. <https://doi.org/10.1038/nn.2831>
- Golomb, J. D. (2015). Divided spatial attention and feature-mixing errors. *Attention, Perception, & Psychophysics*, 77(8), 2562–2569. <https://doi.org/10.3758/s13414-015-0951-0>
- Huang, J., & Sekuler, R. (2010a). Attention protects the fidelity of visual memory: Behavioral and electrophysiological evidence. *Journal of Neuroscience*, 30(40), 13461–13471.
- Huang, J., & Sekuler, R. (2010b). Distortions in recall from visual memory: Two classes of attractors at work. *Journal of Vision*, 10(2), 1–27.
- Kiyonaga, A., & Egner, T. (2016). Center-surround inhibition in working memory. *Current Biology : CB*, 26(1), 64–68. <https://doi.org/10.1016/j.cub.2015.11.013>
- Matsukura, M., Luck, S. J., & Vecera, S. P. (2007). Attention effects during visual short-term memory maintenance: Protection or prioritization? *Perception & Psychophysics*, 69(8), 1422–1434. <https://doi.org/10.3758/BF03192957>
- Nelder, J. A., & Mead, R. (1965). A simplex method for function minimization. *The Computer Journal*, 7(4), 308–313.
- Pratte, M. S., Park, Y. E., Rademaker, R. L., & Tong, F. (2017). Accounting for stimulus-specific variation in precision reveals a discrete capacity limit in visual working memory. *Journal of Experimental Psychology: Human Perception and Performance*, 43(1), 6–17. <https://doi.org/10.1037/xhp0000302>
- 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 of the United States of America*, 109(22), 8780–8785. <https://doi.org/10.1073/pnas.1117465109>
- Wei, X.-X., & Stocker, A. A. (2015). A Bayesian observer model constrained by efficient coding can explain “anti-Bayesian” percepts. *Nature Neuroscience*, 18(10), 1509–1517. <https://doi.org/10.1038/nn.4105>
- Zhang, W., & Luck, S. J. (2008). Discrete fixed-resolution representations in visual working memory. *Nature*, 453(7192), 233–235. <https://doi.org/10.1038/nature06860>

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