BRIEF REPORT

The reliability of retro-cues determines the fate of noncued visual working memory representations

Eren Gunseli · Dirk van Moorselaar · Martijn Meeter ·
Christian N. L. Olivers

Abstract Retrospectively cueing an item retained in visual working memory during maintenance is known to improve its retention. However, studies have provided conflicting results regarding the costs of such retro-cues for the noncued items, leading to different theories on the mechanisms behind visual working memory maintenance and retro-cueing. Here we tested an alternative explanation of the conflicting results regarding retro-cue costs—namely, that they are caused at least partly by differences in retro-cue reliability. We manipulated the ratio of valid-cue trials to invalid-cue trials within blocks. We used a continuous-report procedure that allowed fitting a model that provided recall probability and precision estimates for the memory representations. Reconciling previous contradictory findings, benefits for valid cues were observed in all conditions, but invalid cueing costs were found only when the retro-cue had a high reliability (i.e., was 80 % valid), but not when it had a lower reliability (i.e., 50 % valid). This was found for both the recall probability and the precision of visual working memory representations. Our results suggest that the cognitive mechanisms underlying retro-cue effects are strategically adjusted by participants, depending on the perceived retro-cue reliability.

Keywords Retro-cue · Visual working memory · Attention · Strategy · Reliability

Visual working memory (VWM) is the cognitive system in which a limited amount of visual information can be briefly maintained and manipulated. Attention interacts with many stages of VWM processing, including encoding (Posner, 1980; Schmidt, Vogel, Woodman, & Luck, 2002; Vogel, Luck, & Shapiro, 1998), maintenance (Awh & Jonides, 2001; Awh, Jonides, & Reuter-Lorenz, 1998; Munneke, Heslenfeld, & Theeuwes, 2010), and retrieval (Theeuwes, Kramer, & Irwin, 2011). One way to look at this interaction is through the use of retro-cues. These are typically spatial cues presented during maintenance that point out one of the memory items, which then becomes particularly likely to be tested. It has been shown that such retro-cues result in improved memory performance, relative to trials without a retro-cue (Griffin & Nobre, 2003; Landman, Spekreijse, & Lamme, 2003; Lepsien, Griffin, Devlin, & Nobre, 2005). This retro-cue benefit has been claimed to reflect (a) the reallocation of attentional resources within memory, resulting in the protection of the cued representation against decay and interference (the *protection* hypothesis; Makovski & Jiang, 2007; Makovski, Sussman, & Jiang, 2008; Matsukura, Luck, & Vecera, 2007; Pertzov, Bays, Joseph, & Husain, 2013; van Moorselaar, Gunseli, Theeuwes, & Olivers, under review); (b) removing noncued items from memory, therefore presumably reducing the interitem interference and competition for resources (the *removal* hypothesis; Kuo, Stokes, & Nobre, 2012; Souza, Rerko, & Oberauer, 2014; Williams & Woodman, 2012); (c) carrying the cued item to a more robust or "prioritized" state during maintenance, without altering noncued items (*prioritization during maintenance*; Myers, Walther, Wallis, Stokes, & Nobre, 2014; Rerko & Oberauer, 2013; Souza et al., 2014); or (d) prioritizing the cued representation during retrieval, without affecting maintenance per se (*prioritization during retrieval*; Astle, Summerfield, Griffin, & Nobre, 2012; Nobre, Griffin, & Rao, 2008).

E. Gunseli (✉) · D. van Moorselaar · M. Meeter · C. N. L. Olivers
Department of Cognitive Psychology, VU University, Van der
Boechorststraat 1, 1081 BT Amsterdam, The Netherlands
e-mail: e.gunseli@vu.nl

Although all these hypotheses predict a benefit for the cued representation, they differ in their assumptions regarding the costs for noncued representations. The protection and removal hypotheses predict that retro-cueing benefits for the cued representation should be accompanied by costs for noncued representations, because they involve the reallocation of resources away from noncued items and toward the cued one. The prioritization-during-maintenance and prioritization-during-retrieval hypotheses, on the other hand, predict no such costs, because they explain the cue benefits through a change in the *status* of the cued item, without any change for noncued items. Previous studies using retro-cues have so far produced conflicting results. Some have observed costs in recognition or recall performance when a noncued representation was probed (Matsukura et al., 2007; Pertzov et al., 2013), whereas others have observed no such costs in these so-called *invalid* trials (Landman et al., 2003; Lepsien & Nobre, 2007; Rerko & Oberauer, 2013). The studies that did not observe any invalidity costs used a double-cueing paradigm in which an invalid cue was followed by a valid cue. This may have made participants more hesitant to drop an item after the first cue. However, Matsukura et al. observed a cost of invalid retro-cueing while using a double-cueing paradigm. Therefore, we believe that double-cueing itself cannot account for the inconsistency across findings regarding invalidity costs. Yet again, other studies have observed costs only in reaction times and not in accuracy, which has been interpreted as support for the prioritization-during-retrieval hypothesis (Astle et al., 2012). The lack of costs for noncued items in recognition accuracy has gained further theoretical significance because it has been taken as evidence for the idea that VWM maintenance does not require any active rehearsal via attention (Rerko & Oberauer, 2013; for a similar argument, see Hollingworth & Maxcey-Richard, 2013). Thus, knowing whether retro-cues result in costs for noncued representations is theoretically important for understanding the mechanisms behind VWM maintenance, as well as those behind retro-cueing.

Although we cannot, nor do we wish to, exclude the possibility of different factors playing a role, in the present study we investigated a factor that may at least partially explain the inconsistency in cue-related costs—namely, the reliability of the cue. A similar argument has been made by Williams and Woodman (2012) in the context of directed forgetting cues. Cue reliability can be operationally defined as the ratio between valid and invalid trials.[1] Typically, studies that have failed to reveal a cost in invalid trials have had relatively low cue reliabilities (50 % valid in Landman et al., 2003, and Lepsien & Nobre, 2007; 66.6 % valid in Rerko &

Oberauer, 2013) relative to those in studies that have revealed a cost (80 % valid in Astle et al., 2012; 75 % valid in Matsukura et al., 2007; 70 % valid in Pertzov et al., 2013). It may thus be the case that when a cue has a high reliability (e.g., a high valid-to-invalid trial ratio), participants devote most of their attentional resources to the cued representation (i.e., protection) and remove the noncued items from memory (i.e., removal), since there is very little chance of being tested on them. On the other hand, when a cue has a low reliability (e.g., a low valid-to-invalid trial ratio), participants may keep on maintaining noncued representations in anticipation of potentially being tested on them. In this case, they may merely prioritize the cued item during maintenance and/or retrieval, without costs to the noncued representations. This straightforward hypothesis might account for the inconsistencies in the literature regarding the costs of invalid retro-cues.

We tested the cue reliability hypothesis by manipulating the validity of the retro-cue. Participants were asked to remember the orientations of four bars and then to recall the orientation of one probe bar. On some trials, during the maintenance interval, a probabilistic retro-cue was presented: It pointed to the subsequently probed item in 80 % of trials on some blocks (80 % validity), and in 50 % of trials on other blocks (50 % validity). On the remaining, *invalid* trials, it pointed to one of the items that was *not* subsequently probed. Participants were informed about these validity ratios before each block. Furthermore, rather than employing the often-used change-detection/recognition task, in which observers can only provide a discrete same/different judgment, we used a continuous-recall procedure that we theorized would provide a more sensitive measure of the maintenance of VWM representations, by providing a measure of the degree of quality of recall performance rather than reducing it to a binary decision (Bays, Catalao, & Husain, 2009; Wilken & Ma, 2004; Zhang & Luck, 2008). This measure also enabled us to fit a model that estimated the recall probabilities of the target and nontarget (i.e., nonprobed) representations, and also the precision of memory. To foreshadow the main findings, retro-cues improved the recall probability and precision of the target. Importantly, both the benefits of valid retro-cues and the costs of invalid retro-cues were greater when the cue was highly reliable (i.e., 80 % valid, in comparison to 50 % valid), to the extent that an invalidity cost was absent for both probability and precision in the low-reliability condition.

Method

Experimental procedures

Twenty-two healthy volunteers participated in the experiment for course credit or monetary compensation. For 12 of the participants, we also took electroencephalography (EEG)

---

[1] Note that some studies have defined validity ratios as being calculated as the number of trials valid/(valid + neutral + invalid), instead of as valid/(valid + invalid). In the present study we used the latter formula, since we did not expect trials without a cue to affect the subjective evaluation of the reliability of the cue.

recordings for another study. Their behavioral performance was no different from the rest of the respondents. Two participants were excluded from analysis due to low performance (see the Analysis section). The study was conducted in accordance with the Declaration of Helsinki and was approved by the faculty's Ethics Committee. Written informed consent was obtained.

The procedure is shown in Fig. 1. The memory display consisted of four black oriented bars (2.08° × 0.25° visual angle) located equidistantly on an imaginary circle of radius 3.50°, and was presented for 350 ms. The orientation of each bar was chosen at random, with the restriction that bars within the same trial differed by at least 10°. The test display was presented 1,550 ms after the offset of the memory display (or 1,650 ms, for the 12 participants for whom EEG was also recorded). The test display contained a randomly oriented bar and a cue pointing to the location of the probe representation, both presented at the center of the screen. This probe cue was the same as the fixation circle, except that a quarter (90°) of it was filled white. Participants were asked to indicate the precise orientation of the bar at the probed location by rotating the probe bar using the mouse. After a mouse response was made, the correct orientation was indicated by a central white bar for 100 ms. The intertrial interval was 800 ms for the 10 non-EEG

participants, and was a jittered interval between 1,200 and 1,600 ms for the other 12.

On retro-cue trials, after the memory display, a maintenance interval was presented of either 550 ms (for 10 participants) or 650 ms (for the 12 EEG participants), followed by the presentation of the retro-cue display for 100 ms. The retro-cue was the same as the probe cue, except that the fill color was either red, 27.08 cd/m^2, or green, 24.10 cd/m^2, depending on the reliability condition (order counterbalanced). For the initial practice phase, during which the cue was 100 % valid, the retro-cue fill color was orange (53.46 cd/m^2). Following the retro-cue, there was a second maintenance interval of 900 ms. In no-cue trials, the black fixation circle remained on the screen during the whole maintenance interval, without any changes to it. The timings of the test display were matched for the retro-cue and no-cue trials.

In the high-reliability condition, the cue was 80 % valid, and in the low-reliability condition, it was 50 % valid. The reliability conditions were blocked. In each experimental block, 25 % of the trials were neutral—that is, no cue was presented during the trial. The different validity conditions (i.e., valid, neutral, and invalid) were randomly intermixed within each block. Before each reliability condition, participants were informed about the validity ratio of the cues (which
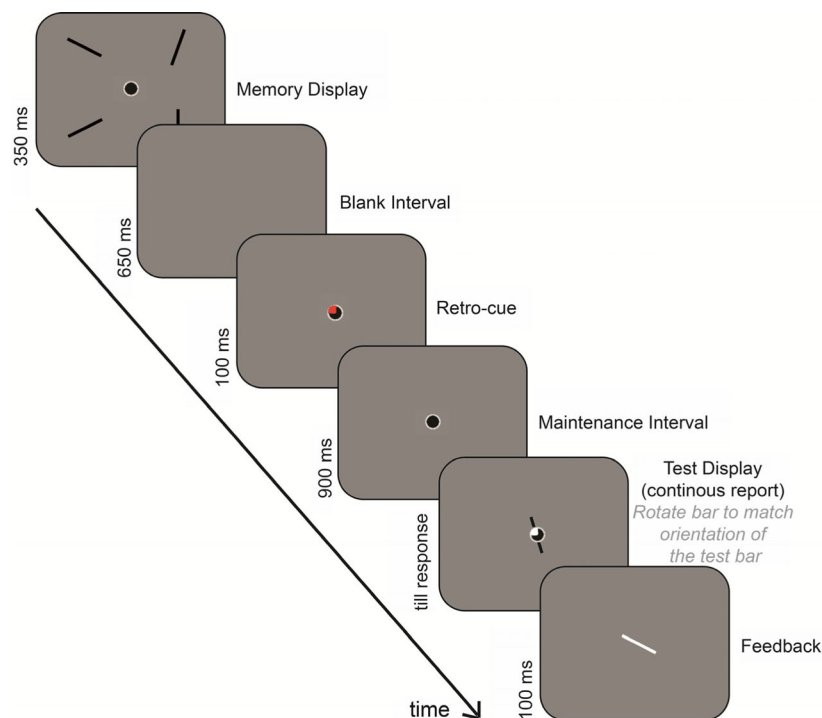
Fig. 1 The experimental procedure in the present experiment. The retro-cue was a fixation circle with a quarter filled with either red or green to point one of the memory representations. Similarly, the test probe was indicated by a white quarter filling. In this example, participants need to report the orientation of the bar presented at the top-left corner, and the retro-cue is valid because it points toward that same position. There were also trials on which the retro-cue would invalidly point to a bar that was not going to be tested. In neutral (i.e., no-cue) trials, the fixation dot remained on the screen during the retro-cue duration. During test, the participants had to rotate the orientation of the bar to match that in their memory by using the mouse.

was also indicated by the color of the cue), and they performed a practice session of 25 trials to get used to this particular validity ratio. Moreover, at the beginning of the experiment, an initial practice session with a 100 %-valid cue was presented, containing 20 trials (25 for the 12 EEG participants), to make the participants familiar with the cue. In total, there were 560 trials (600 for the 12 EEG participants). In order to have a reasonable number of invalid-cue trials, more blocks were presented of the high-reliable cue condition than of the low-reliable cue condition. Respectively for the valid, neutral, and invalid trials, 216, 90, and 54 trials contained an 80 %-valid cue, and 75, 75, and 50 trials (90, 90, and 60 for 12 of the participants) contained a 50 %-valid cue. The main constraint was to have at least 50 trials per condition, to allow for a reliable model fit (see www.paulbays.com/code/JV10/index. php). At the end of each block, participants received feedback on their block average and grand average memory deviation values.

Analysis

Deviation scores on the memory test were calculated as the average difference (i.e., error) between the original orientation of the probed memory bar and the orientation of the response. The precision was calculated, per condition, as the inverse of the standard deviation of the error in the participants' responses (Bays et al., 2009). The deviation scores were entered into a model to calculate the probabilities of recall for the target and nontarget VWM representations (Bays et al., 2009). Two participants were excluded from further analysis due to low performance: One of these had a target recall probability barely above chance level (i.e., above the chance of reporting any of the four orientations) in the 50 %-valid no-cue condition (a fitted recall probability of .26 and an average deviation of 37.1 deg), and the other had a recall probability of .51 (and an average deviation of 34.8 deg) in the 80 %-valid no-cue condition, which was almost the same as the probability of reporting one of the nonprobed items (i.e., .49) for this participant in this condition. Moreover, these target recall probabilities were 2.5 times the standard deviation lower than the overall mean for their given condition. The important results were the same when these participants were included. The raw deviation, target recall probability, nontarget recall probability, and precision for each condition were entered into a repeated measures analysis of variance (ANOVA) with the factors of Reliability (80 % valid vs. 50 % valid) and Validity (valid, neutral, and invalid). Contingent on a significant Reliability × Validity interaction, these were followed up by separate ANOVAs testing for validity benefits (i.e., the difference between neutral and valid trials) and invalidity costs (i.e., the difference between invalid and neutral trials). Where necessary, p values were adjusted for sphericity violations using the Greenhouse–Geisser epsilon correction to the degrees of freedom (Jennings & Wood, 1976).

To test whether the validity benefits and invalidity costs were different from zero, one-sample t tests were used.

Results

We first tested whether the two reliability conditions differed in how participants performed on neutral trials. Any differences on the neutral trials would have suggested that altering the cue reliability changed the way that participants approached the whole task. For all measures described below (i.e., raw deviation, precision, recall probability for the target, and recall probability for nontargets), the performance on neutral trials did not differ between the 80 %-valid and 50 %-valid conditions (all ts < 1.00, ps > .330).

Raw deviations

Next, we looked at the effects of cue validity and reliability on raw deviations from the target orientation. Figure 2A shows the distribution of errors for each condition, where each data point represents the frequency of errors for bins of 15 deg of deviations (Pertzov & Husain, 2014). We found no main effect of reliability on the deviations, $F(1, 19) = 2.10$, $p = .163$, $\eta_p^2 = .10$, but there was an effect of validity, $F(2, 38) = 10.23$, $p < .001$, $\eta_p^2 = .35$. Importantly, a Reliability × Validity interaction was also visible, $F(2, 38) = 26.15$, $p < .001$, $\eta_p^2 = .58$. Planned comparisons showed that both the validity benefit (i.e., the difference between valid and neutral trials), $t(19) = 2.32$, $p = .032$, and the invalidity cost (i.e., the difference between invalid and neutral trials), $t(19) = 2.64$, $p = .023$, were larger for the 80 %-valid condition than for the 50 %-valid condition. Both the validity benefit and the invalidity cost were present (i.e., significantly different from zero) in both reliability conditions (ts > 2.37, ps < .028).

Precision

The average precision in each condition is shown in Fig. 2B. Again, we observed a main effect of validity on precision, $F(2, 38) = 74.54$, $p < .001$, $\eta_p^2 = .80$; no effect of reliability, $F(1, 19) = 0.04$, $p = .838$, $\eta_p^2 = .01$; and a Reliability × Validity interaction, $F(2, 38) = 16.84$, $p < .001$, $\eta_p^2 = .47$. Both the validity benefit, $t(19) = 2.31$, $p = .032$, and the invalidity cost, $t(19) = 3.68$, $p = .002$, were greater in the 80 %-valid than in the 50 %-valid condition. The invalidity cost was significant for the 80 %-valid condition, $t(19) = 4.67$, $p < .001$, but not for the 50 %-valid condition, $t(19) = 0.56$, $p = .580$. The validity benefit was greater than zero in both conditions, ts > 5.35, ps < .001.
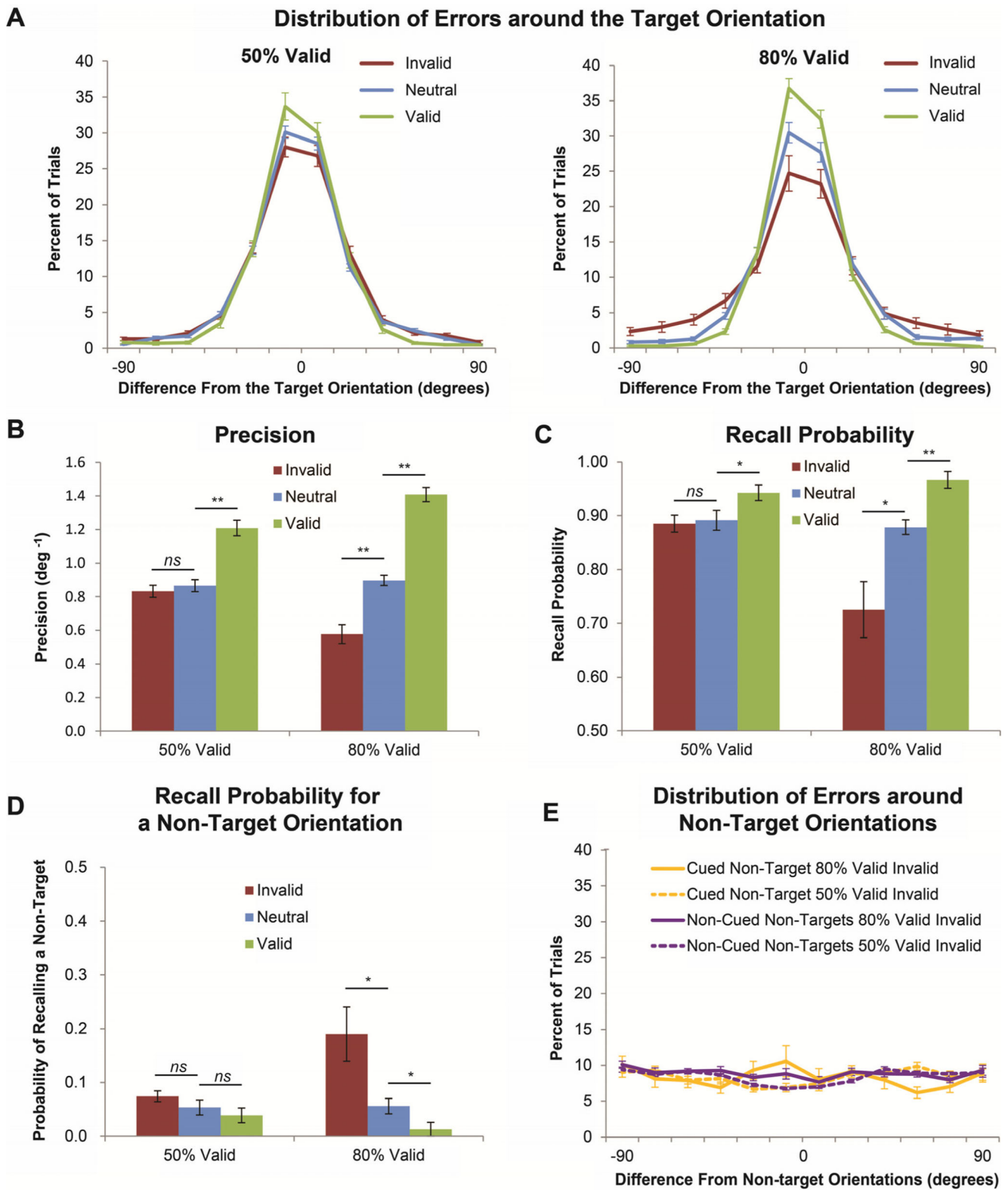
Fig. 2 (A) Distribution of errors relative to the target (i.e., probed) orientation for the 50 %-valid (left panel) and 80 %-valid (right panel) conditions. (B–D) Precision for the target (B), recall probability estimate for the target (C), and recall probability estimate for a nontarget (D) in each condition. The invalid, neutral, and valid trials are shown in different colors, given in the legend. Above the bars, *ns*, *, and ** represent $p >$ .05, $p < .05$, and $p < .005$, respectively. (E) Distribution of errors on invalid trials relative to the nontarget orientations. The error bars represent standard mean errors for the standardized data (i.e., corrected for between-subjects variance; Cousineau, 2005).

Recall probability for the target

The average recall probability in each condition is shown in Fig. 2C. Main effects of reliability, $F(1, 19) = 5.06$, $p = .036$, $\eta_p^2 = .21$, and of validity, $F(2, 38) = 13.75$, $p < .001$, $\eta_p^2 = .42$, on the probabilities were apparent, as well as a Reliability × Validity interaction, $F(2, 38) = 5.91$, $p = .018$, $\eta_p^2 = .24$. Planned comparisons showed that the validity benefit did not differ between the 80 %-valid and the 50 %-valid conditions, $t(19) = 1.41$, $p = .174$, whereas the invalidity cost was larger for the 80 %-valid condition than for the 50 %-valid condition, $t(19) = 2.12$, $p = .047$. The validity benefit was present (i.e., significantly different from zero) for both conditions, $ts > 2.16$, $ps < .044$, whereas the invalidity cost was present in the 80 %-valid condition, $t(19) = 2.45$, $p = .024$, but not in the 50 %-valid condition, $t(19) = 0.26$, $p = .795$.

Recall probability for nontargets

The average probability of recalling a nonprobed item in each condition is shown in Fig. 2D. Here we found a main effect of validity, $F(2, 38) = 7.09$, $p = .013$, $\eta_p^2 = .27$, and none of reliability on this recall probability, $F(1, 19) = 1.61$, $p = .220$, $\eta_p^2 = .08$. Again, a Reliability × Validity interaction emerged, $F(2, 38) = 4.66$, $p = .039$, $\eta_p^2 = .20$. Planned comparisons showed that the validity benefits (in terms of a lower likelihood of recalling a nonprobed item on valid than on neutral trials) were not different for the 80 %-valid and 50 %-valid conditions, $t(19) = 0.12$, $p = .904$, even though the effect was significant, in post-hoc tests, only in the 80 %-valid condition, $t(19) = 2.15$, $p = .045$, and not in the 50 %-valid condition, $t(19) = 1.08$, $p = .294$. The invalidity cost, meanwhile, was higher for the 80 %-valid condition than for the 50 %-valid condition, $t(19) = 2.09$, $p = .051$. The probability of reporting a nonprobed item was greater for invalid than for neutral trials only in the 80 %-valid condition, $t(19) = 2.36$, $p = .029$, but not in the 50 %-valid condition, $t(19) = 0.99$, $p = .336$.

In order to test whether the high probability of reporting a nontarget item was driven by reporting the cued nontarget or any of the (noncued) nontargets, we compared the error distributions around the orientations of both the cued and the noncued nontargets (see Fig. 2E). The distribution of responses around the cued nontarget on 80 %-valid trials was somewhat steeper than those for the cued nontarget on 50 %-valid trials and for the noncued nontargets in both reliability conditions (although the difference in the percentages of errors at –7.5 and 7.5 deg did not reach significance, $ts < 1.73$, $ps > .100$). This leaves open the possibility, although it is statistically not supported, that the higher nontarget recall probability on 80 %-invalid trials in comparison to the other conditions was due to recalling the cued nontarget on a greater proportion of trials rather than to recalling any other nontarget.

Discussion

The findings support the idea that the degree of retro-cue effects on recall performance depends on the reliability of the cue. The cost of invalid cueing was minor for raw deviations, and altogether absent for precision and recall probability estimates when the cue was relatively unreliable (i.e., 50 % valid), whereas there was still a clear benefit for valid cues. When the cue was more reliable (80 % valid), the benefits were larger, and now costs were also present. Furthermore, on invalid trials, the likelihood of mistakenly reporting a nonprobed item during test was higher when the cue was more reliable. These results suggest that how participants implement the retro-cue to the memory task is, at least partly, under strategic control: When the cue has low reliability, observers prioritize the cued item for maintenance and/or retrieval without letting go of the noncued items (prioritization during maintenance and prioritization during retrieval), probably in anticipation of the still quite likely event of being tested on one of the noncued items. As a result, invalid cueing costs are at most minor. In contrast, when the cue is highly reliable, in addition to prioritization, attentional and/or memory resources are disengaged from the noncued items during maintenance (protection and removal), which leads to a high invalidity cost when a noncued item is probed. Retro-cue effects thus seem to be in line with either the prioritization-during-maintenance or prioritization-during-retrieval hypotheses when cue validity is low, but in line with the protection or removal hypotheses when cue validity is high.

Cue reliability may not be the only contributing factor in determining invalidity costs. For example, Astle et al. (2012) found that invalid cues had a cost on recognition accuracy only when memory set size exceeded the VWM capacity limit (i.e., eight), but not for set sizes within the VWM capacity limit (i.e., two and four; but see van Moorselaar, Olivers, Theeuwes, Lamme, & Sligte, 2014), despite the fact that their cue was 80 % valid. Using the same set size and the same cue reliability, in the present study we observed a cost of invalid cueing. Our study and that of Astle et al. differed in the test used to measure memory performance. We believe that the continuous-report procedure used in the present study is a more sensitive memory measure, and therefore might reveal differences in performance that are less likely to be detected with the discrete same/different judgment task that was used by Astle et al., because it provides a measure of how good the response is for each trial instead of reducing the response to a binary decision (Wilken & Ma, 2004). Consistent with this claim, in the present study, the effects of retro-cueing were more pronounced for precision than for the recall probability of items. Nevertheless, we cannot exclude the possibility that invalidity costs might be smaller, although not completely absent, for smaller set sizes, even with a continuous-recall measure, since the possibility of being tested on a particular

noncued item is higher, and also maintenance is less demanding for smaller set sizes. Both of these factors make it less beneficial to redistribute attentional/memory resources when the set size is small.

Notwithstanding the role of set size, our findings suggest that some of the inconsistency in results regarding invalidity costs is due to differences in the reliability of the retro-cue (for a similar argument for directed forgetting cues, see Williams & Woodman, 2012). Thus, the absence of an invalidity cost in Rerko and Oberauer (2013) may merely reflect a lack of attentional redistribution, due to the low reliability of the retro-cue, rather than the absence of a role of attention in VWM maintenance. Our results suggest that attentional redistribution is performed mostly for highly reliable cues (as in the 80 %-valid cue condition in the present study) and that without being attended, VWM representations are vulnerable to interference and/or decay—consistent with earlier claims (Astle et al., 2012; Makovski & Jiang, 2007; Makovski et al., 2008; Matsukura et al., 2007; Pertzov et al., 2013; van Moorselaar, Gunseli, et al., 2014). Another possibility is that noncued items are actively removed from memory when cues are highly reliable, since this would also result in significant invalidity costs (Kuo et al., 2012; Souza et al., 2014; Williams & Woodman, 2012). Considering that previous research has provided support for both mechanisms, we believe that both may occur. On some trials, noncued items may be actively removed from memory, whereas on other trials they are attended less, and therefore are more vulnerable to interference. The important conclusion that we want to make is that either mechanism is more likely to be implemented when the cue is highly reliable.

Regardless of these exact mechanisms, our findings point to a dissociation between how retro-cues affect the cued and noncued items: While the cued item is attended, the noncued items may be unattended, but not necessarily dropped (Rerko & Oberauer, 2013). In other words, whether a noncued item is maintained or not may be a separate decision than whether it is attended or not. Such a dissociation is consistent with several models promoting a distinction between memory items that are in the current focus ("template") and other VWM representations that are held prospectively, or "on reserve" (LaRocque, Lewis-Peacock, Drysdale, Oberauer, & Postle, 2013; LaRocque, Lewis-Peacock, & Postle, 2014; Oberauer, 2002; Olivers, Peters, Houtkamp, & Roelfsema, 2011; Rerko & Oberauer, 2013; van Moorselaar, Olivers, et al., 2014; van Moorselaar, Theeuwes, & Olivers, 2014; Zokaei, Manohar, Husain, & Feredoes, 2014). These two types of VWM representations may have different mechanisms of maintenance, which operate more or less independently: (1) Task-relevant (here cued) representations are carried into a prioritized template status (which may also prioritize them for retrieval) regardless of the cue reliability, since there is little to lose by doing so. In line with this idea, Berryhill, Richmond, Shay, and Olson (2012)

demonstrated the presence of a validity benefit even when the cue was informative only on a minority of trials. (2) Currently irrelevant noncued items are held via a more passive accessory storage, and observers may decide to remove these depending on the perceived reliability of the cue—that is, depending on whether they see a potential future use for them.

In short, the present results show that how retro-cues affect recall performance depends on the reliability of the cues. When they are highly reliable, retro-cues resulted in major invalidity costs and larger validity benefits (as compared to low-reliable retro-cues, which resulted in minor invalidity costs and smaller validity benefits). Thus, cue reliability will have to be considered before drawing any conclusions from research using probabilistic retro-cues.

References

Astle, D. E., Summerfield, J., Griffin, I., & Nobre, A. C. (2012). Orienting attention to locations in mental representations. *Attention, Perception, & Psychophysics, 74,* 146–162. doi:10.3758/s13414-011-0218-3

Awh, E., & Jonides, J. (2001). Overlapping mechanisms of attention and spatial working memory. *Trends in Cognitive Sciences, 5,* 119–126. doi:10.1016/S1364-6613(00)01593-X

Awh, E., Jonides, J., & Reuter-Lorenz, P. A. (1998). Rehearsal in spatial working memory. *Journal of Experimental Psychology: Human Perception and Performance, 24,* 780–790. doi:10.1037/0096-1523.24.3.780

Bays, P. M., Catalao, R. F. G., & Husain, M. (2009). The precision of visual working memory is set by allocation of a shared resource. *Journal of Vision, 9*(10), 7:1–11. doi:10.1167/9.10.7

Berryhill, M. E., Richmond, L. L., Shay, C. S., & Olson, I. R. (2012). Shifting attention among working memory representations: Testing cue type, awareness, and strategic control. *Quarterly Journal of Experimental Psychology, 65,* 426–438.

Cousineau, D. (2005). Confidence intervals in within-subject designs: A simpler solution to Loftus and Masson's method. *Tutorials in Quantitative Methods for Psychology, 1,* 42–45.

Griffin, I. C., & Nobre, A. C. (2003). Orienting attention to locations in internal representations. *Journal of Cognitive Neuroscience, 15,* 1176–1194. doi:10.1162/089892903322598139

Hollingworth, A., & Maxcey-Richard, A. M. (2013). Selective maintenance in visual working memory does not require sustained visual attention. *Journal of Experimental Psychology: Human Perception and Performance, 39,* 1047–1058. doi:10.1037/a0030238

Jennings, J. R., & Wood, C. C. (1976). The epsilon-adjustment procedure for repeated-measures analyses of variance. *Psychophysiology, 13,* 277–278.

Kuo, B.-C., Stokes, M. G., & Nobre, A. C. (2012). Attention modulates maintenance of representations in visual short-term memory. *Journal of Cognitive Neuroscience, 24,* 51–60. doi:10.1162/jocn_a_00087

Landman, R., Spekreijse, H., & Lamme, V. A. F. (2003). Large capacity storage of integrated objects before change blindness. *Vision Research, 43,* 149–164. doi:10.1016/S0042-6989(02)00402-9

LaRocque, J. J., Lewis-Peacock, J. A., Drysdale, A. T., Oberauer, K., & Postle, B. R. (2013). Decoding attended information in short-term memory: An EEG study. *Journal of Cognitive Neuroscience, 25,* 127–142. doi:10.1162/jocn_a_00305

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

Lepsien, J., Griffin, I. C., Devlin, J. T., & Nobre, A. C. (2005). Directing spatial attention in mental representations: Interactions between attentional orienting and working-memory load. *NeuroImage, 26,* 733–743. doi:10.1016/j.neuroimage.2005.02.026

Lepsien, J., & Nobre, A. C. (2007). Attentional modulation of object representations in working memory. *Cerebral Cortex, 17,* 2072–2083. doi:10.1093/cercor/bhl116

Makovski, T., & Jiang, Y. V. (2007). Distributing versus focusing attention in visual short-term memory. *Psychonomic Bulletin & Review, 14,* 1072–1078. doi:10.3758/BF03193093

Makovski, T., Sussman, R., & Jiang, Y. V. (2008). Orienting attention in visual working memory reduces interference from memory probes. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 34,* 369–380. doi:10.1037/0278-7393.34.2.369

Matsukura, M., Luck, S. J., & Vecera, S. P. (2007). Attention effects during visual short-term memory maintenance: Protection or prioritization? *Perception & Psychophysics, 69,* 1422–1434. doi:10.3758/BF03192957

Munneke, J., Heslenfeld, D. J., & Theeuwes, J. (2010). Spatial working memory effects in early visual cortex. *Brain and Cognition, 72,* 368–377. doi:10.1016/j.bandc.2009.11.001

Myers, N. E., Walther, L., Wallis, G., Stokes, M. G., & Nobre, A. C. (2014). Temporal dynamics of attention during encoding versus maintenance of working memory: Complementary views from event-related potentials and alpha-band oscillations. *Journal of Cognitive Neuroscience.* doi:10.1162/jocn_a_00727

Nobre, A. C., Griffin, I. C., & Rao, A. (2008). Spatial attention can bias search in visual short-term memory. *Frontiers in Human Neuroscience, 1*(4), 1–9. doi:10.3389/neuro.09.004.2007

Oberauer, K. (2002). Access to information in working memory: Exploring the focus of attention. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 28,* 411–421. doi:10.1037/0278-7393.28.3.411

Olivers, C. N. L., Peters, J., Houtkamp, R., & Roelfsema, P. R. (2011). Different states in visual working memory: When it guides attention and when it does not. *Trends in Cognitive Sciences, 15,* 327–334. doi:10.1016/j.tics.2011.05.004

Pertzov, Y., Bays, P. M., Joseph, S., & Husain, M. (2013). Rapid forgetting prevented by retrospective attention cues. *Journal of Experimental Psychology: Human Perception and Performance, 39,* 1224–1231. doi:10.1037/a0030947

Pertzov, Y., & Husain, M. (2014). The privileged role of location in visual working memory. *Attention, Perception, & Psychophysics, 76,* 1914–1924. doi:10.3758/s13414-013-0541-y

Posner, M. I. (1980). Orienting of attention. *Quarterly Journal of Experimental Psychology, 32,* 3–25. doi:10.1080/00335558008248231

Rerko, L., & Oberauer, K. (2013). Focused, unfocused, and defocused information in working memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 39,* 1075–1096. doi:10.1037/a0031172

Schmidt, B. K., Vogel, E. K., Woodman, G. F., & Luck, S. J. (2002). Voluntary and automatic attentional control of visual working memory. *Perception & Psychophysics, 64,* 754–763. doi:10.3758/BF03194742

Souza, A. S., Rerko, L., & Oberauer, K. (2014). Unloading and reloading working memory: Attending to one item frees capacity. *Journal of Experimental Psychology: Human Perception and Performance, 40,* 1237–1256. doi:10.1037/a0036331

Theeuwes, J., Kramer, A. F., & Irwin, D. E. (2011). Attention on our mind: The role of spatial attention in visual working memory. *Acta Psychologica, 137,* 248–251. doi:10.1016/j.actpsy.2010.06.011

van Moorselaar, D., Gunseli, E., Theeuwes, J., & Olivers, C. N. L. (2014). The time course of protecting a visual working memory representation from visual interference. *Frontiers in Human Neuroscience, 8.* doi:10.3389/fnhum.2014.01053

van Moorselaar, D., Theeuwes, J., & Olivers, C. N. L. (2014c). In competition for the attentional template: Can multiple items within visual working memory guide attention? *Journal of Experimental Psychology: Human Perception and Performance, 40,* 1450–1464. doi:10.1037/a0036229

Vogel, E. K., Luck, S. J., & Shapiro, K. L. (1998). Electrophysiological evidence for a postperceptual locus of suppression during the attentional blink. *Journal of Experimental Psychology: Human Perception and Performance, 24,* 1656–1674. doi:10.1037/0096-1523.24.6.1656

Wilken, P., & Ma, W. J. (2004). A detection theory account of change detection. *Journal of Vision, 4*(12), 1120–1135. doi:10.1167/4.12.11

Williams, M., & Woodman, G. F. (2012). Directed forgetting and directed remembering in visual working memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 38,* 1206–1220. doi:10.1037/a0027389

Zhang, W., & Luck, S. J. (2008). Discrete fixed-resolution representations in visual working memory. *Nature, 453,* 233–235. doi:10.1038/nature06860

Zokaei, N., Manohar, S., Husain, M., & Feredoes, E. (2014). Causal evidence for a privileged working memory state in early visual cortex. *Journal of Neuroscience, 34,* 158–162. doi:10.1523/JNEUROSCI. 2899-13.2014