Author(s): Fu, Xueying; Ye, Chaoxiong; Hu, Zhonghua; Li, Ziyuan; Liang, Tengfei; Liu, Qiang

Title: The impact of retro-cue validity on working memory representation : Evidence from electroencephalograms

Year: 2022

Version: Accepted version (Final draft)

Please cite the original version:

Fu, X., Ye, C., Hu, Z., Li, Z., Liang, T., & Liu, Q. (2022). The impact of retro-cue validity on working memory representation : Evidence from electroencephalograms. Biological psychology, 170, Article 108320. https://doi.org/10.1016/j.biopsycho.2022.108320

The impact of retro-cue validity on working memory representation: Evidence from electroencephalograms

Xueying Fu [1,2+], Chaoxiong Ye [1,3,4+], Zhonghua Hu[1], Ziyuan Li [5], Tengfei Liang [5],

Qiang Liu [1,5] *

[1] Institute of Brain and Psychological Sciences, Sichuan Normal University, 610000, Chengdu, China.

[2] Department of Cognitive Neuroscience, Faculty of Psychology and Neuroscience, Maastricht University, EV Maastricht, 6229, the Netherlands.

[3] Department of Psychology, University of Jyvaskyla, 40014, Jyvaskyla, Finland.

[4] Center for Machine Vision and Signal Analysis, University of Oulu, 90014, Oulu, Finland.

[5] Research Center of Brain and Cognitive Neuroscience, Liaoning Normal University, 116029, Dalian, China.

+ Xueying Fu and Chaoxiong Ye contributed equally to this work and should be considered as co-first authors.

* Correspondence should be addressed to Qiang Liu, Institute of Brain and Psychological Sciences, Sichuan Normal University, Chengdu, 610000, China. Telephone: +8613332220573. E-mail: lq780614@163.com.

1. Introduction

Visual working memory (VWM) plays an essential role in cognitive processing and performance. VWM has been proposed as a cognitive system that temporarily stores and manipulates visual information to meet the needs of ongoing cognitive tasks (Luck & Vogel, 1997, 2013). In recent years, a growing body of research has explored the mechanisms of VWM, and researchers now suggest that VWM is a flexible, dynamic process rather than a fixed one (Christophel, Iamshchinina, Yan, Allefeld, & Haynes, 2018; Christophel, Klink, Spitzer, Roelfsema, & Haynes, 2017; Ma, Husain, & Bays, 2014; Myers, Chekroud, Stokes, & Nobre, 2018; Wolff, Jochim, Akyurek, & Stokes, 2017; Ye et al., 2017; Ye et al., 2020; Ye et al., 2019). However, the VWM capacity is extremely limited, so the visual system is often exposed to demanding tasks that exceed its limits. Thus, mechanisms of selective attention are needed to control access to VWM and to prioritize the existing VWM representations for behavioral output.

Attentional prioritization in VWM has been extensively studied using retro-cues (Souza & Oberauer, 2016). In a typical retro-cue experiment (Griffin & Nobre, 2003; Landman, Spekreijse, & Lamme, 2003), participants are asked to remember a memory array for subsequent reporting. During the interval between the memory array and the probe array, a retro-cue is presented to indicate which item of the memory array is most likely to be tested. The effect of the retro-cue on VWM performance is called the retro-cue effect and includes retro-cue benefits and retro-cue costs. A retro-cue benefit refers to improved memory performance resulting from a valid retro-cue condition (i.e., the location of the to-be-tested item is indicated), and this can be calculated by the difference in behavioral performance between the valid retro-cue condition and the neutral retro-cue condition. Conversely, a retro-cue cost refers to impaired memory performance resulting from an invalid retro-cue (i.e., the location of an item that will not be tested is indicated), and this can be calculated as the difference in behavioral performance between an invalid retro-cue condition and a neutral retro-cue condition.

Recent studies have investigated the underlying mechanisms of the retro-cue effect by examining whether this effect is modulated by the validity of the retro-cue (calculated as $\frac{\text{number of trials of valid cue condition}}{\text{number of trials of valid cue condition} + \text{number of trials of invalid cue condition}}$) (Günseli et al., 2019; Günseli, van Moorselaar, Meeter, & Olivers, 2015). For example, by manipulating the retro-cue validity, one behavioral study showed that retro-cue benefits were clearly observed regardless of the retro-cue validity (Günseli et al., 2015). However, the retro-cue costs could only be unambiguously identified when the retro-cue had high validity (i.e., 80% validity). The retro-cue costs were minor for raw deviations (a metric for memory quality) and absent for memory precision and memory probability when the retro-cue had low validity (i.e., 50% validity). Based on the behavioral results, Günseli et al. (2015) suggested that when the cue was relatively unreliable, the participants would prioritize the cued representation for maintenance without dropping the non-cued representations. By contrast, when the cue was highly reliable, in addition to prioritizing, the participants would drop the non-cued representations during maintenance, thereby incurring obvious retro-cue costs when a non-cued item was tested. That is, the mechanisms underlying retro-cue effects can be strategically (or automatically) adjusted by the participants; therefore, the removal of the non-cued representations from VWM will depend on the expected validity of the retro-cue.

A potential problem arises with the behavioral results of the retro-cue studies because many additional innate processing stages, such as encoding, retrieval, and decision-making, may also possibly affect the behavioral results of the VWM task. These extra processing stages could potentially contribute to a behavioral outcome, thereby corrupting the measurement of VWM storage (Keshvari, van den Berg, & Ma, 2013). Therefore, behavioral results may not provide sufficiently strong evidence to confirm that retro-cue validity affects selective attention and storage during VWM maintenance before the test probe. These behavioral results complicate the unambiguous detection of the retro-cue effect in VWM.

One technique for tracking the VWM process online without potential contamination by other processes is to use electroencephalograms (EEGs), and several researchers have previously used EEGs to investigate the retro-cue effect (Goddertz, Klatt, Mertes, & Schneider, 2018; Kuo, Stokes, & Nobre, 2012; Poch, Valdivia, Capilla, Hinojosa, & Campo, 2018; Schneider, Barth, Getzmann, & Wascher, 2017). For example, lateralized alpha powers (8–14 Hz) and contralateral delay activity (CDA) have been used as indicators of attention and VWM maintenance. The lateralized alpha power is widely accepted as being able to track the locus of covert visuospatial attention (Bacigalupo & Luck, 2019; Ikkai, Dandekar, & Curtis, 2016; Klatt, Getzmann, Wascher, & Schneider, 2018; Poch, Capilla, Hinojosa, & Campo, 2017; Poch et al., 2018; Sauseng et al., 2005; Thut, Nietzel, Brandt, & Pascual-Leone, 2006;

Worden, Foxe, Wang, & Simpson, 2000). The alpha power over the parietal-occipital electrodes in the hemisphere contralateral to the attended item is reduced relative to the ipsilateral electrodes, both during and after the perception, within VWM. Conversely, CDA is an accepted metric for tracking VWM storage (Feldmann-Wustefeld, Vogel, & Awh, 2018; Luria, Balaban, Awh, & Vogel, 2016; Vogel & Machizawa, 2004; Vogel, McCollough, & Machizawa, 2005). CDA appears as a sustained negative waveform over the parietal-occipital electrodes in the hemisphere contralateral to the remembered stimuli. The CDA amplitude is thought to track the number of stored items in an online maintenance state within VWM.

A follow-up study on retro-cue validity by Günseli et al. (2019) used EEGs to measure the VWM process during maintenance after retro-cue onset. The authors manipulated the validity of retro-cues and asked participants to conduct a continuous report task. Their memory array contained three different orientations: one presented on the vertical midline and the other two presented left and right from the center. Two retro-cue validity conditions (80% validity and 50% validity) were included in their study. The behavioral results showed that the retro-cue effect (error on the invalid cue trials − error on the valid trials) was larger under the high-validity condition than under the low-validity condition. The EEG results showed obvious lateralized alpha powers under both the high-validity and low-validity conditions, but no difference was evident between the validity conditions at the beginning of the task. However, at about 700 ms from the onset of the retro-cue, the lateralized alpha power under the low-validity condition returned to baseline, resulting in a significant difference between the low-validity and high-validity conditions in the latter part of the interval period.

These results suggested that participants paid attention to the cued item when the retro-cue was presented under both low-validity and high-validity conditions, but they sustained attentional prioritization for a longer period in the high-validity condition than in the low-validity condition. The researchers also noted the emergence of an obvious CDA early after the retro-cue in the high-validity state, whereas CDA in the low-validity state appeared only later in the trial, just before the onset of the probe array. Importantly, this difference in the CDA amplitude under conditions of high and low validity generally became apparent early in the interval period, rather than later. The study by Günseli et al. (2019) requires that the participants remember the stimuli of both the left and right hemifields simultaneously; therefore, CDA could serve as an index for measuring asymmetrical maintenance in VWM. That is, CDA should not be found when items are encoded/maintained equally in both hemifields, whereas obvious CDA should appear if participants drop the non-cued representations from online memory (unequal memory load in two hemifields). Thus, the CDA results reported by Günseli et al. (2019) suggested that, although the process time course may differ, the non-cued representations were eventually dropped from VWM under both high-validity and low-validity conditions. This result seems inconsistent with the results of the same group's earlier study (Günseli et al. 2015), which had suggested that participants would continue to maintain the non-cued items when the retro-cue validity is low.

We propose two potential explanations for the evidence suggesting the prolonged maintenance of non-cued items under the low-validity conditions, as observed by Günseli et al. (2019). One is that those researchers used a continuous report task to measure VWM performance. Consequently, their participants needed to memorize, with high precision, orientations that were considered more complex than simple materials (e.g., colors) (Hao, Becker, Ye, Liu, & Liu, 2018; Stevanovski & Jolicoeur, 2011; Vogel, Woodman, & Luck, 2006). The imposed task required that the participants report the target item as precisely as possible, thereby encouraging the participants to concentrate all their VWM resources on one item to maintain its representation with high precision. Thus, the participants had a strong motivation to drop the non-cued representations from VWM.

The second explanation may be that Günseli et al. (2019) used a 50% valid retro-cue as the low-validity condition. Compared to the chance level of 33% for memorizing three items and detecting one of them, the participants could obtain the benefit of an extra 17% chance under the low-validity condition if they used a retro-cue to remove the non-cued representations from VWM. The choice of whether to maintain a non-cued representation in VWM may represent a strategic control (or a result of implicit statistical learning); however, individual differences exist in the control of these strategies adopted by participants under the 50% validity condition. Quite possibly, even under a 50% validity condition, the participants could use the same strategy (resource allocation mechanism) that they use under the high-validity condition (80% validity). This would lead to the eventual removal of the non-cued representations from VWM under the 50% validity condition. In that case, the retro-cue effect would be caused by both retro-cue benefits (i.e., the strengthening of the cued representation) and retro-cue costs (i.e., the loss of non-cued representations) under both the low-validity (50%

validity) and the high-validity (80% validity) conditions. Günseli et al. (2019) did not establish a neutral cue condition in their study; consequently, they could not confirm this second possibility because they could not identify whether the retro-cue effect was due to the contribution of retro-cue benefits or retro-cue costs.

The aim of the present study was to test whether retro-cue validity affects the fate of non-cued representations in VWM. We used EEGs to investigate how non-cued representations are handled in VWM under different cue validity conditions. We also used an improved experimental design to minimize the pitfalls apparent in the study by Günseli et al. (2019). In our study, the participants conducted a change-detection task to remember four colored squares that were symmetrically distributed on both the left and right visual fields. We manipulated the validity of the retro-cue and recorded the EEG signals to explore the prolonged selective attention and memory storage process after the onset of the retro-cue. The impact of retro-cue validity was investigated by setting the validity of the retro-cue to 80% valid (the high-validity state) and 20% valid (the low-validity state; this was slightly below the chance level of 25%) across the experimental blocks. We set the cue validity to 20% valid as the low-validity condition because we did not want participants to gain extra performance benefits by allocating additional attention/memory resources to the cued item under the low-validity condition. Therefore, under the low-validity condition, the participants should not have a conscious motivation to allocate more resources to the cued item. On the contrary, they should have a stronger motivation to allocate resources to the non-cued items under the low-validity condition. In this case, if a retro-cue effect is still obvious under the low-validity condition, this would suggest that the retro-cue effect may be partly driven by bottom-up processes.

We also established the cause of the retro-cue effect by setting neutral cue trials to identify the retro-cue benefit (i.e., better performance in valid cue trials than in neutral cue trials) and the retro-cue cost (i.e., worse performance in valid cue trials than in neutral cue trials). Under both the high-validity and low-validity conditions, we used lateralized alpha power to track the prolonged selective attention and CDA to index VWM storage, as described by Günseli et al. (2019). We determined the prolonged selective attention to the cued item by observing whether a sustained lateralized alpha power emerged (i.e., whether a smaller alpha power contralateral to the cued item was evident). For memory, we assumed that because participants needed to encode and maintain the items in both hemifields at the same time, no asymmetry would be apparent in the EEG signal (i.e., no CDA would emerge) if the participants continued to maintain all items in VWM. By contrast, when non-cued items (particularly those from the hemifield opposite the cued item) were dropped from memory, CDA would be expected to emerge (i.e., a stronger negativity contralateral to the cued item should be evident). We anticipated that the retro-cue would redirect selective attention to a cued item under both the low-validity and high-validity conditions; however, the participants would maintain the non-cued representations during the interval under the low-validity condition while dropping them from VWM under the high-validity condition. Thus, we expected to observe lateralized alpha power after the retro-cue appeared under both the low-validity and high-validity conditions. We also expected to observe CDA only under the high-validity condition and not under the low-validity condition.

2. Methods

2.1. Participants

Adequate power for the t-test comparison was ensured by a priori determination of the sample size by a power analysis based on the predicted effect size using G∗Power 3.1.9.2 (Faul, Erdfelder, Lang, & Buchner, 2007). According to the study by Berryhill, Richmond, Shay, and Olson (2012), the difference between different cue conditions has a medium effect size (e.g., Cohen's d = 0.49, for Experiment 3, Ignore mixed) for accuracy. Thus, we assumed a medium effect size (Cohen's d = 0.50) for our experimental design. For a statistical power of $(1 - \beta) = 0.80$ and a significance level of 0.05, the suggested total sample size was 34 participants. The suggested sample size in our study is slightly larger than the sample size used in previous similar studies (i.e., 22 participants in the study by Günseli et al. (2015); 30 participants in the study by Günseli et al. (2019)).

In total, 38 students (16 males and 22 females) volunteered to take part in our experiment for compensation. All participants were healthy and right-handed, with normal or corrected-to-normal vision. No individuals reported achromatopsia, anomalous trichromatism, or psychiatric disorders. Two of these 38 participants were excluded from further analysis because of a ceiling effect in their behavioral performance (accuracy close to 100% in the neutral cue), and another two were excluded because of extreme artifacts in their EEG data (the number of available trials was less than 50 on either side of each validity). Ultimately, data from 34 participants (19 females and 15 males) were used for the final statistical analyses (mean age: 20.59 ± 1.76 years; range 18 to 25 years). All participants provided written consent before enrollment in the study and received a monetary reward (25 CNY per hour). All procedures in our study were conducted in accordance with the Declaration of Helsinki (2008) and were approved by the Ethics Committee of Liaoning Normal University.

2.2. Materials

The experiment was programmed using E-prime software (E-prime 2.0, Psychology Software Tools, Inc.). The color stimuli were four colored squares (each 1° × 1°), randomly chosen from red (255,0,0), green (0,255,0), blue (0,0,255), yellow (255,255,0), white (255,255,255), magenta (255,0,255), purple (128,0,128), orange (255,125,0), and turquoise (64,224,208). All stimuli were displayed on a 19-inch CRT monitor (60 Hz) on a gray (128,128,128) background at a viewing distance of 70 cm.

2.3. Procedure

The participants were asked to perform a change-detection task with a retro-cue. Two different retro-cue types were used: an informative cue that pointed to the location of one memory item and a neutral cue that pointed to all four locations of the memory items. The validity (80% validity or 20% validity) of the informative retro-cue was manipulated under different blocks for each participant.

The trial structures are depicted in Figure 1. Each trial began with a central fixation that appeared for 200 ms. A memory array of colored squares was then presented at the corners of an invisible square (5° × 5°) for 100 ms. The participants were instructed to memorize the colors of the four colored squares. After 500 ms had elapsed from the offset of the memory array (first delay), an informative retro-cue or a neutral cue was presented at the center for a duration of 100 ms. The cue was then followed by the rest of the retention interval (second delay), with a duration of 800 ms. After that interval, the participants were asked to indicate whether the probe stimulus was identical to the memory item (50%) or if the color had changed to a new color that had not appeared in the presented memory array at the corresponding location (50%). The next trial started at 800–1400 ms after the response. Before the task, participants were instructed to stare at the fixation, to minimize eye blinks, and to respond as accurately as possible. The retro-cue type was selected at random during each trial, based on the validity condition.

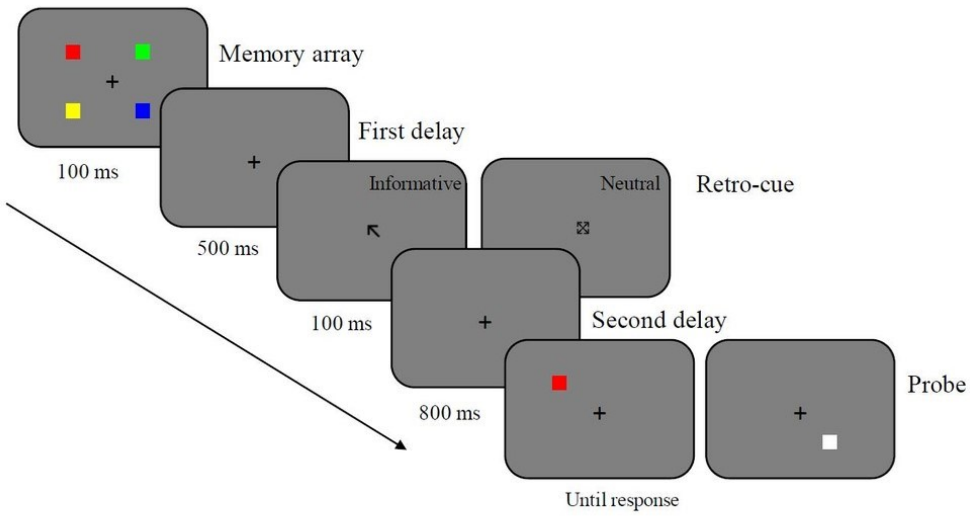

Fig. 1. *Retro-cue experimental design. At the beginning of each trial, a new memory array, including four color items, was presented symmetrically. The participants were requested to remember the items and complete the change-detection task after a retro-cue. These*

retro-cues came in two forms: informative (66.7%) and neutral (33.3%). The informative retro-cue was divided into a valid cue and an invalid cue. The whole experiment consisted of a high-validity (e.g., 80% validity) condition and a low-validity (20% validity) condition in separate blocks.

For each participant, the first half of the experiment consisted of one validity condition and the other half consisted of the other validity condition. The high-validity and low-validity conditions were blocked, and the order of the blocks was counterbalanced across participants. A 2 × 3 repeated-measures design, including the within-subject factors of cue validity (high vs. low) and cue type (neutral vs. valid vs. invalid), was employed in our experiment. We manipulated the retro-cue validity (high-validity vs. low-validity) in two different blocks. The total number of trials of each block was the same, but the ratio of valid cue trials (i.e., pointing to the location of the to-be-tested item) and invalid cue trials (i.e., pointing to the location of a not-to-be-tested item) differed under different validity blocks.

Overall, 360 trials were run for each validity condition in the formal experiment. Each block was divided into six mini-blocks of 60 trials each, with a break of at least 30 s between mini-blocks and 2 min between blocks. Each mini-block consisted of 20 neutral trials and 40 informative retro-cue trials (32 valid cue trials and 8 invalid cue trials in the high-validity condition; 8 valid cue trials and 32 invalid cue trials in the low-validity condition). We ensured that participants were familiar with the formal experiment by informing them of the validity of the retro-cues (80% validity for the high-validity condition and 20% validity for the low-validity condition) and having them perform a practice block of 50 trials (including 10 neutral trials) before each validity block with the same validity as the block. In the practice block, feedback about whether the response was correct or wrong was given after each trial. In the formal experiment, feedback was provided regarding the overall accuracy during each break. The participant was required to show an accuracy in the practice block that exceeded 75% for the experiment to continue. The experiment took approximately 1 h.

2.4. EEG recording

The EEG data were recorded from a 64-electrode cap (BioSemi ActiveTwo, BioSemi Inc., Amsterdam, the Netherlands) using the International 10/20 System. Two additional electrodes on both sides around the vertex (Cz) were used as the online reference and ground electrodes. Electrodes were also placed on the right and left mastoid as off-line references. F7/F8 were placed at the left and right outer corners, 1 cm away from the eyes, to monitor horizontal eye movements (HEOG). Fpz was used to monitor vertical eye movements (VEOG). EEG signals were amplified and digitized at a sampling rate of 512 Hz, with 24-bit resolution and no online filter. Electrode impedances were kept below 5 kΩ.

2.5. Data analysis

A significance level of $p < 0.05$ was used for all tests. A repeated-measures ANOVA was applied to test the effects of validity and cue type on accuracy. The assumption of sphericity was assessed by Mauchly's tests, and the Greenhouse-Geisser correction was applied to adjust the degrees of freedom for violations of sphericity. Partial eta squared (η^2_p) measures were used for effect size estimations for the ANOVAs. The t-tests were conducted using a bootstrapping method (SPSS Statistics Version 23; 10,000 permutations with 95% confidence intervals). Cohen's d was used as an estimator of the effect size for the t-tests. We also used JASP software (Version 0.16, JASP Team, 2021) to conduct Bayes factor analyses (Bayesian t-test) to show whether the results favored the alternative hypothesis or the null hypothesis (Rouder, Speckman, Sun, Morey, & Iverson, 2009). The default priors in JASP were used (Schmalz, Biurrun Manresa, & Zhang, 2021). The Bayes factor (BF_{10}) provides an odds ratio for the alternative/null hypotheses (values < 1 favor the null hypothesis and values > 1 favor the alternative hypothesis). For example, a BF_{10} of 4 would indicate that the alternative hypothesis is 4 times more likely than the null hypothesis to be correct, while a BF_{10} of 0.2 would indicate that the null hypothesis is 5 times more likely than the alternative hypothesis to be correct. The results of CDA and lateralized alpha power were corrected for multiple comparisons using false discovery rate (FDR) correction (Benjamini & Hochberg, 1995) at a statistical threshold of $p < .05$ (MATLAB 2015b, MathWorks, Inc., Natick, MA). We also calculated the two-tailed Pearson's correlation coefficients between the ERP indicators (CDA, lateralized alpha power) and behavioral indicators (accuracy of each cue type, retro-cue benefit index, and retro-cue cost index under high-validity condition or low-validity condition).

2.5.1. Behavioral data analysis

The accuracy of three different cue types (valid, invalid, and neutral) in two validity blocks was calculated to assess memory performance. A repeated-measures ANOVA with validity condition (high-validity, low-validity) and cue type (valid, neutral, invalid) as within-subject factors was conducted for accuracy. The interaction effects found in ANOVAs were followed up using paired-samples *t*-tests (two-tailed) conducted for pairwise comparison of the different cue types under both the high-validity and the low-validity conditions using Bonferroni correction.

2.5.2. EEG data preprocessing.

We analyzed the EEG data from trials with neutral, valid, and invalid cues during VWM maintenance. Off-line EEG data were processed in MATLAB (2015b, MathWorks, Inc., Natick, MA) using the EEGLAB toolbox 14.1.2 (Delorme & Makeig, 2004) and scripts. As in the study by Günseli et al. (2019), the scalp EEG was band-pass filtered (cutoff frequencies: 0.01 Hz and 40 Hz) and re-referenced off-line to the average of the left and right mastoids. Continuous EEG data were epochal, from −500 ms to 2000 ms around the memory array onset in each trial. Trials in which the EEG amplitude exceeded ± 50 μV at HEOG (F7/F8) and ± 75 μV at PO7/PO8 and VEOG (Fpz) during the 0 to 1500 ms interval (time-locked to the memory array onset) were deemed to contain artifacts and were rejected. Additional blinks and eye or head movements were rejected based on visual inspection. Subsequently, the epoch was based on the direction of the retro-cue (pointed to the left or right side). Two participants were excluded from the final sample because they had fewer than 50 trials on either side in each condition after artifact rejection. On average, we retained 91 ± 19 left-side epochs and 91 ± 18 right-side epochs per participant under high-validity conditions and 92 ± 17 left-side epochs and 94 ± 16 right-side epochs per participant under low-validity conditions for further analysis. We also investigated the effects of eye movements on the EEG measures of interest and determined that eye movements were unlikely to spuriously generate the EEG dynamics (CDA and lateralized alpha power) effect we observed in the present study. More details are provided in the Supplementary Materials.

In the main text, we have mainly focused on the results of valid cue trials and invalid cue trials. Upon further analysis, we reported only the EEG data of trials with informative retro-cues in the Results section. The EEG results of neutral cue trials under different validity conditions (which is not the focus of this paper) can be found in the Supplementary Materials.

2.5.3. Analysis of CDA amplitudes

As with some recent CDA studies (Feldmann-Wustefeld & Vogel, 2019; Feldmann-Wustefeld et al., 2018; Hakim, Feldmann-Wüstefeld, Awh, & Vogel, 2020), we chose the PO7/PO8 electrodes for the analyses of CDA amplitudes, using a 200 ms prior to memory array onset as the baseline (-200 to 0 ms, time-locked to the memory array onset). However, since CDA has been shown to be present and large at many electrode pairs (McCollough, Machizawa, & Vogel, 2007), many studies have used multiple electrode pairs for data analysis when investigating the CDA component (Günseli et al., 2019; Günseli, Meeter, & Olivers, 2014; Günseli, Olivers, & Meeter, 2014; Gao, Xu, et al., 2011; Gao, Yin, Xu, Shui, & Shen, 2011; Hao et al., 2018; Ikkai, McCollough, & Vogel, 2010; Liang et al., 2020; Ngiam, Adam, Quirk, Vogel, & Awh, 2021; Peterson, Gozenman, Arciniega, & Berryhill, 2015; Wang, Rajsic, & Woodman, 2019; Ye et al., 2018; Ye, Zhang, Liu, Li, & Liu, 2014). We also chose the multiple parietal-occipital electrode pairs (P5/P6, P7/P8, and PO7/PO8) to reanalyze our data in the Supplementary Materials; the result pattern using multiple electrode pairs was highly consistent with the result using the PO7/PO8 electrodes. More details are provided in the Supplementary Materials. Hence, we only reported the results using the PO7/PO8 electrodes in the main text.

For CDA, the contralateral waveforms were computed as the average of the activity recorded at the left hemisphere electrode sites when the retro-cues (including valid and invalid cues) pointed to the right side of the memory array and the average of the activity recorded from the right hemisphere electrode sites when they pointed to the left side. The ipsilateral waveforms were computed by averaging the left and right hemisphere sites when the cues pointed to the left or the right side of the memory array, respectively. CDA was defined by subtracting the ipsilateral activity from the contralateral activity.

We assumed that other memory and cognitive processes were present before the retro-cue appeared. Thus, we chose to use the memory array onset with baseline correction during the EEG analysis, as was done in some previous studies (Goddertz et al., 2018; Schneider et al., 2017), and we focused on the stage of VWM maintenance after the onset of the retro-cue. During the period before the probe array appeared, the participants did not know whether the retro-cue in the trial was valid. The validity of the retro-cue in each trial was determined by the probe array. The use of this design meant that we did not need to analyze the valid retro-cue trials or the invalid retro-cue trials in the EEG data. Instead, we analyzed all trials with an informative retro-cue (both valid and invalid) and compared the differences in the EEG results between the informative retro-cue trials under the high-validity and low-validity conditions.

Previous studies have shown that CDA can be observed 300–400 ms after retro-cue onset and that it persists throughout maintenance (Günseli et al., 2019; Vogel & Machizawa, 2004; Williams & Woodman, 2012). This established a time window of interest between 900 and 1500 ms (300–900 ms after retro-cue onset) in this study. The amplitudes of the different waves (cue contralateral − cue ipsilateral) at each time point over the whole time window (0–1500 ms) were calculated under high-validity or low-validity conditions. For the within-condition testing, we conducted a two-tailed one-sample *t*-test against zero (Groppe, Urbach, & Kutas, 2011) with false discovery rate (FDR) correction (Benjamini & Hochberg, 1995) at each time point to test for the presence of a significant lateralized component. The mean amplitudes of CDA across the time window of interest (900–1500 ms) under the high-validity and low-validity conditions were compared with the zero value by a one-sample t-test.

Subsequently, for the between-condition testing, the t-tests (two-tailed, FDR correction) were applied at each time point to compare the amplitude of the different waves under the high-validity and low-validity conditions. The mean amplitudes of CDA across the time window of interest (900–1500 ms) under the high-validity and low-validity conditions were compared with the paired-samples t-test.

2.5.4. Analysis of lateralized alpha power

Similar to the analysis of the CDA component, we chose the PO7/PO8 electrodes for the analyses of lateralized alpha power. As shown in the Supplementary Materials, we also chose parietal-occipital electrode pairs (P5/P6, P7/P8, and PO7/PO8) to reanalyze our data; the result pattern using multiple electrode pairs was highly consistent with the result using the PO7/PO8 electrodes. We maintained consistency in the EEG analyses by only reporting the results obtained using the PO7/PO8 electrodes here in the main text. We conducted a time–frequency analysis of the alpha-band power by convoluting the trials that were the same as the CDA analysis with a complex Morlet wavelet transform (width: seven cycles, from 1 to 30 Hz in 1 Hz increments). We used −300 to 0 ms relative to the memory array onset as the baseline (L. Zhang, Peng, Zhang, & Hu, 2013; Z. G. Zhang, Hu, Hung, Mouraux, & Iannetti, 2012). As with the CDA analysis, the lateralized alpha power was calculated as the difference between the contralateral and ipsilateral dB-normalized power values. The power values were averaged across the alpha band (8–14 Hz), and we selected a time window of interest of 900–1500 ms (300–900 ms after retro-cue onset). The power of the lateralized alpha band at each time point was calculated under both conditions and analyzed by a one-sample *t*-test (two-tailed) with FDR correction over the time window. The power difference of the lateralized alpha band between the high-validity and low-validity conditions was compared by paired-samples t-tests (two-tailed, FDR correction). The mean power of the lateralized alpha band during the time window under the high-validity and low-validity conditions was compared with the zero value by a one-sample t-test, whereas the difference between the two validity conditions was compared with the paired-samples t-test.

We also replicated the analysis of previous studies (L. Zhang et al., 2013; Z. G. Zhang et al., 2012) by providing an extra analysis of CDA and the lateralized alpha power using the time window prior to retro-cue onset for baseline correction (see Supplementary Material). We generally observed highly consistent results for both CDA and the lateralized alpha power, regardless of the baseline correction analysis employed.

2.5.5. Correlation analysis

We assessed the ability of the EEG signal (CDA/lateralized alpha power) to predict behavioral performance and whether CDA and lateralized alpha power influenced each other or were independent during VWM. We first calculated the mean CDA amplitude and the mean lateralized alpha power under each validity condition in the time window of interest (900–1500 ms) for each participant. We then applied Pearson correlation (two-tailed) analysis to investigate the relation between behavioral indicators (the accuracy of the valid cue and invalid cue types, the retro-cue benefit index [valid – neutral], and the retro-cue cost index [neutral – invalid]) and EEG indicators (CDA amplitude and lateralized alpha power). We also calculated the Pearson's correlation (two-tailed) between CDA amplitude and the lateralized alpha power.

3 Results and discussion

3.1. Behavioral results

The accuracies in each cue type for the high-validity and low-validity conditions are shown in Figure 2. The ANOVA showed a significant main effect of cue type (F (1.586, 52.345) = 146.765, p <.001, η_p^2 = .816) and a significant interaction between the cue type by validity condition (F (2, 66) =18.237, p <.001, η_p^2 = .356), but we found no significant main effect of validity condition (F (1,33) = 0.617, p =.438, η_p^2 = .018). The paired-samples t-tests revealed significantly higher accuracy for the valid cue trials (0.935 ± 0.043 for high validity; 0.902 ± 0.068 for low validity) than for the neutral cue trials (0.792 ± 0.071 for high validity; 0.794 ± 0.079 for low validity) under both the high-validity condition (t (33) = 14.200, p <.001, 95% CI [0.123, 0.164], d = 2.434, BF_{10} > 10000) and the low-validity condition (t (33) = 7.837, p <.001, 95% CI [0.080, 0.136], d = 1.344, BF_{10} > 10000). The accuracy was significantly greater for the neutral cue trials than for the invalid cue trials (0.722 ± 0.082 for high validity; 0.770 ± 0.084 for low validity) under both the high-validity condition (t (33) = 6.246, p <.001, 95% CI [0.047, 0.092], d = 1.071, BF_{10} > 10000) and the low-validity condition (t (33) = 2.636, p =.038, 95% CI [0.005, 0.041], d = 0.452, BF_{10} = 3.529). By contrast, the accuracy was significantly greater for the valid cue trials under the high-validity condition than under the low-validity condition (t (33) = 3.850, p = .0015, 95% CI [0.016, 0.051], d = 0.660, BF_{10} = 57.824), whereas the accuracy was significantly lower for the invalid cue trials under the high validity condition than under the low validity condition (t (33) = 3.535, p = .0037, 95% CI [-0.075, -0.020], d = 0.606, BF_{10} = 26.585). No significant difference was noted in the accuracy of the neutral cue trials between the high-validity and the low-validity condition (t (33) = 0.211, p = .834, 95% CI [-0.018, 0.015], d = 0.036, BF_{10} = 0.188). These results indicated that significant retro-cue benefits were obtained from the valid cues and significant retro-cue costs were incurred from the invalid cues under both the high-validity and low-validity conditions. The retro-cue benefits and the retro-cue costs were also larger under the high-validity condition than under the low-validity condition.

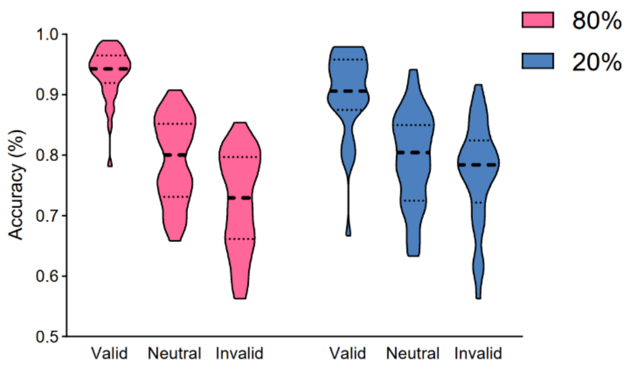

Fig. 2. *Violin plots of the behavioral performance for the high-validity condition (80% validity, pink) and the low-validity condition (20% validity, blue). The dashed line and two dotted lines indicate the median and the two quartiles.*

3.2. CDA results

Figure 3 shows the CDA amplitude results and the grand-averaged difference waveform of the retro-cue trials under the high-validity and low-validity conditions. For the within-condition testing, each time point from 0–1600 ms was corrected against zero using the false discovery rate (FDR) at a statistical threshold of $p < .05$. No significant difference was observed for the waves under either validity condition or between conditions before the retro-cue onset, indicating that no lateralized ERP component was present before the retro-cue under either validity condition. Significant CDA was observed after a retro-cue with a high-validity (898–1600 ms, FDR-corrected of $p < .05$) rather than a low-validity retro-cue. Studying the time window of interest of 900–1500 ms revealed that the mean CDA amplitude under the high-validity condition was significantly different from zero (t (33) = 5.174, p <.001, 95% CI [-1.483, -0.671], d = 0.887, BF_{10} = 1890.109), suggesting that CDA was present under the high-validity condition during the entire retaining period after the retro-cue. However, the mean CDA amplitude did not differ from zero under the low-validity condition (t (33) = 0.362, p = .720, 95% CI [-0.428, 0.299], d = 0.062, BF_{10} = 0.195), suggesting the absence of any obvious CDA component under the low-validity condition.

For the between-condition testing, the FDR-corrected results for the time window from 900–1600 ms also showed a statistically significant difference (shown as the black bar, FDR-corrected of p < .05) in CDA amplitude between the two validity conditions over the time courses of 898–986 ms, 1002–1063 ms, and 1389–1600 ms. The mean CDA amplitude was larger under the high-validity condition than under the low-validity condition (t (33) = 3.748, p = .001, 95% CI [-1.525, -0.471], d = 0.643, BF_{10} = 44.886). These results suggested that the participants maintained the non-cued items during the interval under the low-validity condition but dropped them from VWM under the high-validity condition, as expected.

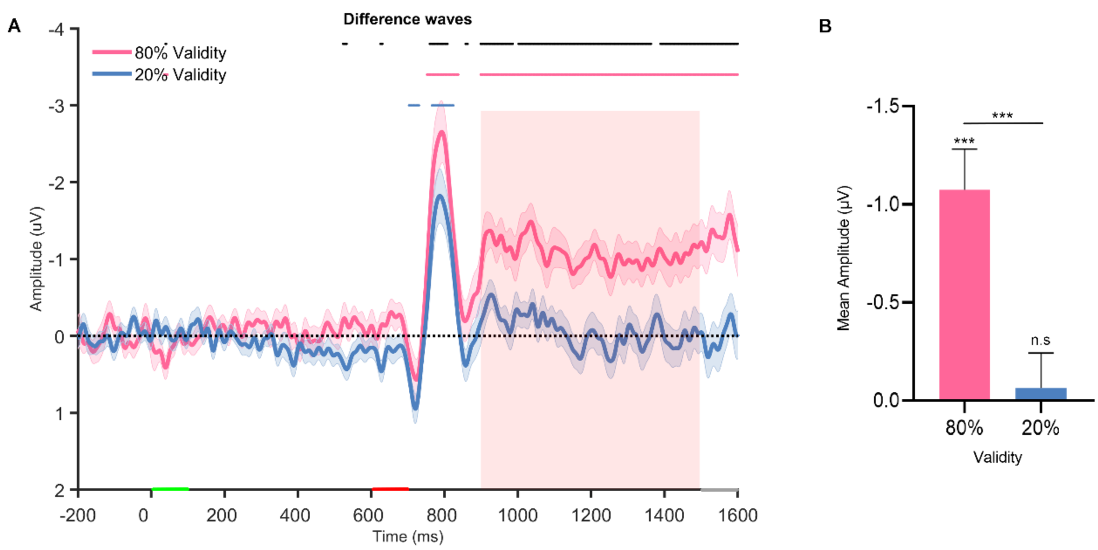

Fig. 3 *Difference waves during the entire time window and the CDA amplitude results. (A) Mean ERP difference waveforms time-locked to the onset of the memory array under the high-validity condition (80% validity) and the low-validity condition (20% validity). The green, red, and gray rectangles on the x-axis show the timing of the memory array (0–100 ms), retro-cue (600–700 ms), and probe (1500–1600 ms), respectively. The red shadow shows the time window of interest (900–1500 ms). The shadow of the curve indicates the standard error of the estimate. The black lines along the tops of the waves indicate a significant difference in amplitude over the entire time course between the 80% validity and 20% validity conditions. The pink lines along the top of the waves indicate an amplitude significantly larger than zero under the 80% validity condition over the entire course. The blue lines along the top of the waves indicate an amplitude significantly larger than zero under the 20% validity condition over the entire time course; (B) The mean CDA amplitude results (900–1500 ms) under the high-validity and the low-validity conditions are displayed as pink and blue bars, respectively (Error bar: SE). ** = p < .01; *** = p < .001; n.s. = not significant.*

3.3. Lateralized alpha power results

Figure 4 illustrates the lateralized alpha power results and the grand-averaged lateralized alpha power of the retro-cue trials under the high-validity and low-validity conditions. For the within-condition testing, the lateralized alpha power was detected over the entire maintenance period after the retro-cue under the high-validity condition (890–1600 ms, one-sample t-test with an FDR-corrected p-value of 0.05). Most time points after the retro-cue appeared under the low-validity condition (1025–1600 ms, one-sample t-test with an FDR-corrected p-value of 0.05). Studying the time window of 900–1500 ms revealed that the mean lateralized alpha power across the time window was significantly different from zero under both the high-validity condition (t (33) = 4.923, p < .001, 95% CI [-2.275, -1.006], d = 0.844, BF_{10} = 957.937) and the low-validity condition (t (33) = 3.316, p = .002, 95% CI [-1.750, -0.419], d = 0.569, BF_{10} = 15.782), suggesting an obvious lateralized alpha power under both the high-validity and low-validity conditions.

For the between-condition testing, no difference in lateralized alpha power was detected between the high-validity and low-validity conditions at any time point (FDR-corrected of p < .05). For the mean lateralized alpha power, no significant difference was detected over the time window of interest (900–1500 ms) between the two validity conditions (t (33) = 1.295, p = .204, 95% CI [-1.359, 0.243], d = 0.222, BF_{10} = 0.395). As expected, these results suggested that retro-cues redirect selective attention to the cued item under both the low-validity and the high-validity conditions.

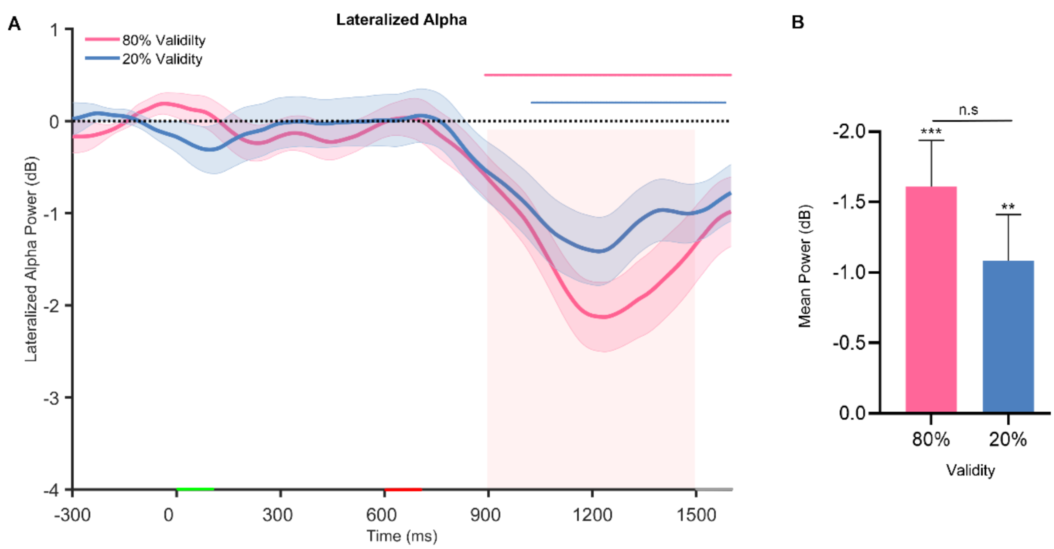

Fig. 4. *The lateralized alpha power waves and their mean power during the time window of interest. (A) Mean lateralized alpha power time-locked to the onset of the memory array for the high-validity condition (80% validity) and the low-validity condition (20% validity). The green, red, and gray rectangles on the x-axis show the timing of the memory array (0–100 ms), retro-cue (600–700 ms), and probe (1500–1600 ms), respectively. The red shadow shows the time window of interest (900–1500 ms). The shadow of the curves is the standard error of the estimate. The black lines along the top of the waves indicate a significant difference in amplitude over the time course between the 80% validity and 20% validity conditions (the absence of a black line indicates no significant difference). The pink lines along the top of the waves indicate that the amplitude under the 80% validity condition is significantly larger than zero over the time course. The blue lines along the top of the waves indicate that the amplitude under the 20% validity condition is significantly larger than zero over the time course. (B) The mean power of the lateralized alpha band (900–1500 ms) under the high-validity and low-validity conditions is displayed as pink and blue bars, respectively (Error bar: SE). ** = $p < .01$; *** = $p < .001$; n.s = not significant.*

3.4. Correlation results

The EEG results showed no correlation between the CDA amplitude and the lateralized alpha power under either validity condition (r (34) = -0.023, p = .899 for high validity; r (34) = 0.115, p = .519 for low validity). Similarly, no significant correlation was detected between the CDA amplitude and the mean accuracies among the valid and invalid cues under either validity condition (high-validity: all $p > .338$; low-validity: all $p > .052$) or between the lateralized alpha power and the mean accuracies under either validity condition (high-validity: all *p* > .318; low-validity: all *p* > .234). The results showed no significant correlation between the retro-cue benefit/cost index and the CDA amplitude (high-validity: all *p* > .510; low-validity: all *p* > .258). Similarly, no significant correlation was detected between the lateralized alpha power and the retro-cue benefit/cost index under either validity condition (high-validity: all *p* > .556; low-validity: all *p* > .466). Taken together, our findings provided no evidence of any significant correlation between the behavioral performance, CDA amplitude, and lateralized alpha power under either the high-validity or the low-validity conditions. That is, the correlation analyses did not show any evidence that the CDA amplitude could be predicted by the alpha-band power or by behavioral performance.

3.5. Exploratory analysis

A visual inspection of the difference waveforms (Fig 3A) suggested that an N2pc component was elicited after the onset of a retro-cue under both validity conditions. An N2pc is usually observed at the posterior electrode on the contralateral side of the target position at 200 ms after the lateralization stimulus onset, and it reflects the spatial attention placed on the target location. Previous work has shown that an N2pc is more negative when elicited by the target item than by the non-target items when multiple items are presented (Eimer, 1996; Luck & Hillyard, 1994a, 1994b; Zhao et al., 2011). A relatively common practice is to interpret the N2pc (180–320 ms) as an index of the deployment of covert lateraled visual attention (Kiss, Van Velzen, & Eimer, 2008) or of the onset of attentional engagement (Zivony, Allon, Luria, & Lamy, 2018). Therefore, based on our visual inspection, and similar to previous studies using CDA (Allon & Luria, 2019; Feldmann-Wustefeld & Vogel, 2019), we conducted exploratory analyses to explore the N2pc components under our different conditions.

The preprocessing and calculation of the amplitudes of the difference waveforms of the N2pc component were conducted essentially as described for the CDA component. In the present study, the FDR-corrected results showed that the N2pc was averaged from the difference wave at the PO7/PO8 electrodes during the 780–880 ms after the memory array onset (180–280 ms after the retro-cue onset) for each condition. One-sample t-tests were applied to detect whether a significant N2pc was elicited, and a paired-samples t-test was applied to detect whether the N2pc amplitude showed a significant difference under the different validity conditions. The correlation between the N2pc and the CDA/lateralized alpha power was also analyzed using the Pearson correlation. Note that we have no a priori assumptions about the N2pc in this research; therefore, these findings should be interpreted with caution.

The N2pc was also measured under the high-validity and low-validity conditions (i.e., 80% validity and 20% validity, respectively) as the difference in the mean amplitude between the ipsilateral and contralateral waveforms recorded at the analyzed electrodes (PO7/PO8) (Feldmann-Wustefeld & Vogel, 2019; Luck & Hillyard, 1994a, 1994b) at 780–880 ms after memory array onset (180–280 ms after the retro-cue onset).

Significant N2pc components were found after the high-validity retro-cue ($M = -1.277 \pm 1.543$; $t\,(33) = 4.822$, $p < .001$, 95% CI [-1.797, -0.770], $d = 0.827$, $BF_{10} = 730.289$) and the low-validity retro-cue (M = -0.674 ± 1.186; t (33) = 3.315, p = .002, 95% CI [-1.103, -0.312], d = 0.568, BF_{10} = 15.750), suggesting that the N2pc component was reliably observed under both validity conditions. The N2pc amplitude was significantly larger under the high-validity condition than under the low-validity condition (t (33) = 2.549, p = .016, 95% CI [-1.046, -0.134], d = 0.437, BF_{10} = 2.960). These results indicated that a retro-cue could redirect the participants' attention to the cued hemifield, regardless of the validity of the retro-cue. By contrast, more attention resources were allocated to cued items after a more reliable retro-cue appeared.

Significant positive correlations were detected between the N2pc amplitude and the CDA amplitude in the 80% validity (r = .536, p = .001) and 20% validity conditions (r = .473, p = .005), whereas no significant correlation was found between the N2pc and the lateralized alpha power under either validity condition (all $p > .055$). We also did not find any significant correlation between the N2pc amplitude and memory recognition performance (all $p > .323$) or between the N2pc amplitude and the retro-cue benefit/retro-cue cost index (all $p > .140$) for either validity condition.

4. General discussion

In the present study, we tested whether the retro-cue validity affects the fate of non-cued representations in VWM. We found significant retro-cue benefits and retro-cue costs for the behavioral results under both the high-validity and low-validity conditions. More importantly, for the EEG results, although the retro-cue could redirect selective attention to the cued hemifield under both low-validity and high-validity conditions, the participants maintained the non-cued items during the interval under the low-validity condition, whereas they dropped them from VWM under the high-validity condition.

4.1. Retro-cues work in both top-down and bottom-up processes

Our behavioral results demonstrate the retro-cue effects occurring under both high-validity and low-validity conditions. A previous study indicated that the retro-cue effect persists even when the cue validity is set at the chance level (Berryhill et al., 2012). Similarly, our results demonstrate a retro-cue benefit for a valid cue even when the cue validity was set below the chance level (20%). This suggests that a retro-cue could automatically guide attention even when the cue was disadvantageous (i.e., it cues an item that is likely to be irrelevant). Our EEG results also support this suggestion. The retro-cue clearly could guide participants' attention under the high-validity condition; however, an obvious lateralized alpha power and the N2pc component were also observed after retro-cue appearance under the low-validity condition, similar to those under the high-validity condition. These results indicate that the retro-cue effect is not fully under optimal strategic control. Under the low-validity condition, the optimal strategy was to allocate similar/fewer resources to the cued item compared to the non-cued items, indicating an imperfect resource allocation mechanism. Our work revealed that retro-cues can be used partly in a bottom-up manner (i.e., stimulus-driven).

A natural question also arises regarding the possibility that the impact of cue validity on the retro-cue effect is also affected by top-down control. Our results showed that both the retro-cue benefit and retro-cue cost were larger under the high-validity condition than under the low-validity condition. That is, the expectation of cue validity had an impact on the degree of the retro-cue effect. In addition, a difference was evident in the mechanisms underlying the retro-cue effect between the high-validity and low-validity conditions. These results suggest that the impact of cue validity on the retro-cue effect was at least partly caused by top-down control.

Two possible top-down control methods could explain our results. One is strategic control. In our experiment, the participants were informed in advance of the cue validity (i.e., 20% or 80%) in each validity condition, and they were allowed sufficient practice before performing the formal experiment. Therefore, the expectation of cue validity can lead the participants to use the retro-cue strategically. The other possible control method is implicit statistical learning. In this case, the participants automatically carried out implicit statistical learning during the experiment and formed an optimal memory mechanism (e.g., whether to keep the non-cued items in VWM) to obtain better performance. Future studies can test these two possibilities by intermixed vs. blocked reliability manipulation or by controlling whether the participants are informed of the cue validity in advance.

4.2. Impact of retro-cue validity on EEG dynamics

We noted discrepancies between the impact of retro-cue validity on CDA and the lateralized alpha power. An obvious CDA component was found under the high-validity condition but was absent under the low-validity condition. This demonstrates that, based on the expectation of retro-cue validity, the participants could store only the cued representation in the online maintenance state within VWM under the high-validity condition, whereas they could store both the cued and the non-cued representations in VWM under the low-validity condition. Thus, our CDA results can be regarded as supplementary EEG evidence for the EEG results reported by Günseli et al. (2015).

For the lateralized alpha power results, although our results and those reported by Günseli et al. (2019) suggest that the participants shifted their attention to the cued item after the retro-cue appeared, regardless of the cue validity, some small differences are still apparent between our results and the results reported by Günseli et al. (2019). In contrast to the findings of Günseli et al. (2019), we found a similar and sustained lateralized alpha power during VWM maintenance after the retro-cue onset and until the onset of the probe array under both the high-validity and low-validity conditions. Günseli et al. (2019) reported that, regardless of retro-cue validity, the lateralized alpha power disappeared before the probe array onset under both the high-validity and low-validity conditions. The difference in the duration of the lateralized alpha power between our study and theirs may reflect the differences in experimental design. In our study, we used a change-detection task, which required participants to allocate their attention to the location of the probe stimulus and then compare it with the memory item after the probe array appeared. By contrast, Günseli et al. (2019) used a continuous report task, which required participants to reallocate their attention to the center of the screen to adjust the probe item after the probe array appeared. In our study, the participants would focus steadily on the cued item's position to complete the task, leading to sustained lateralized alpha power. Conversely, the participants in the study by Günseli et al. (2019) would shift their attention back to the center of the screen after they had allocated attention to the cued item. This would lead to a lateralized alpha power that would emerge only during early maintenance and then vanish. Therefore, the differences observed in the lateralized alpha power results between our study and that by Günseli et al. (2019) could be due to differences in the setting of the probe array.

4.3. Attention and storage are two distinct processes in VWM

In the present study, the lateralized alpha power reflects the selective attention of the participants, while the N2pc component reflects the attention redirection of the participants. However, we did not find any significant correlation between the lateralized alpha power and the N2pc component. This result was in line with the findings of a recent study (Bacigalupo & Luck, 2019), which found that the lateralized alpha power and the N2pc have different time courses and influence mechanisms, suggesting that they reflect a related but distinct attention mechanism. The lateralized alpha power was of greater interest in the present study because of its similar time course to that of the CDA component.

The idea that attention can predict storage in VWM remains controversial. Some studies have shown a correlation between CDA and lateralized alpha power (van Dijk, van der Werf, Mazaheri, Medendorp, & Jensen, 2010), but a growing number of studies now suggest that the neural mechanism of CDA and lateralized alpha power are separated within VWM (Bae & Luck, 2018; Fukuda, Mance, & Vogel, 2015; Günseli et al., 2019; Hakim, Adam, Günseli, Awh, & Vogel, 2019). Lateralized alpha power is considered a good metric for measuring prolonged selective attention, and CDA reliably predicts representative storage levels. Günseli et al. (2019) observed a dissociation between the CDA amplitude and lateralized alpha power with the retro-cues of different reliabilities in a continuous report task. In corroboration with their findings, we failed to observe any correlation between the CDA amplitude and lateralized alpha power, regardless of the cue validity. Especially under the low-validity condition, we detected obvious lateralized alpha power but no CDA component. These results again support the idea that storage and prolonged selective attention in VWM are two distinct processes that can operate differently, depending on the needs of the task.

In addition to indicating a relation between prolonged selective attention and storage, our results showed a positive correlation between the N2pc amplitude and the CDA amplitude, regardless of the cue validity. These results were in line with the findings reported by Salahub and Emrich (2020). Our results demonstrate that an increased likelihood of dropping a non-cued representation in VWM (as indicated by the CDA) is associated with an increased reallocation of attention to the cued item (as indicated by the N2pc). Therefore, although prolonged selective attention and storage are two distinct processes, the process of redirection of attention could predict the process of VWM storage.

4.4. Different mechanisms for the retro-cue effect under different expectations of cue validity

As predicted, the mechanisms underlying the retro-cue effect varied depending on the cue validity. Previous studies have proposed different hypotheses to explain the cause of the retro-cue effect (Souza & Oberauer, 2016). Our results suggest that, under the high-validity condition, the participants redirected their attention to the cued item and allocated VWM resources to it, while reducing or removing memory resources from the non-cued items. Therefore, under the high-validity condition, the participants could gain significant retro-cue benefits from the valid cue. This mechanism could be considered consistent with a removal hypothesis, suggesting that retro-cues can help to reduce memory load by removing non-cued items from VWM, thereby freeing up VWM resources to maintain the cued item (Goddertz et al., 2018; Kuo et al., 2012; Poch et al., 2018; Williams, Hong, Kang, Carlisle, & Woodman, 2013).

By contrast, under the low-validity condition, although the participants in our study could still redirect their selective attention to the cued item, they allocated equal VWM resources to the cued item and to other non-cued items in an attempt to maintain all items to the greatest extent possible. In this case, the participants could still significantly gain retro-cue benefits from the valid cue. These results would appear to refute the removal hypothesis.

The mechanism of the retro-cue effect could instead be interpreted by an attentional strengthening hypothesis, suggesting that attention is redirected to augment the accessibility of cued representations in VWM (Goddertz et al., 2018; Kuo et al., 2012; Poch et al., 2018; Williams et al., 2013). However, this attentional strengthening hypothesis does not specifically state what happens to the non-cued representations. That is, the attentional strengthening hypothesis does not preclude the possibility that an additional mechanism also operates on the non-cued items. Thus, the maintenance of non-cued representations under the low-validity condition does not conflict with the attentional strengthening hypothesis.

The key to testing the removal hypothesis is to investigate whether participants drop the non-cued representations from VWM after the retro-cue appears. The study of Günseli et al. (2019) showed ERP evidence (i.e., the CDA component) supporting the removal hypothesis (i.e., the participants eventually dropped the non-cued representations from VWM) under both high-validity and low-validity conditions. By contrast, we found ERP evidence supporting the removal hypothesis only under the high-validity condition.

The difference between the findings of Günseli et al. (2019) and our results under the low-validity condition may stem from the fact that we used a below-chance level (20% validity) as the low validity to reduce strategy conflicts caused by ambiguous validity (50% validity), thereby amplifying the underlying mechanism of low validity. This is in line with the inference that participants might take a longer time to change the status of non-cued items under a 50% valid condition (i.e., the low-validity condition used by Günseli et al. (2019)).

One point to consider is that we only found EEG evidence supporting or rejecting the removal hypothesis at the group level. By contrast, at the individual level, we found no significant correlation between EEG results and any of the behavioral measures. Thus, our results did not really allow disentanglement of the extent to which the behavioral retro-cue benefit/cost is due to prolonged selective attention (i.e., lateralized alpha power) or to removal of non-cued items from VWM (i.e., a CDA effect). However, the lack of a significant correlation between the EEG and behavioral results could have several explanations. For example, the behavioral results in the change-detection task are affected by the VWM process, but they are also influenced by the decision-making process. The need for extra processing stages could potentially contribute to a behavioral outcome. This may explain why behavioral indicators are less sensitive than EEG indicators (i.e., CDA) for reflecting the representations of VWM storage. In addition, because our experimental design mainly focuses on the results of group-level comparisons, the failure to find a significant correlation may be due to the limited number of participants. Therefore, the findings regarding a relationship between EEG results and the retro-cue benefit/cost should be interpreted with caution.

We also found that although our participants maintained the non-cued representations in VWM under the low-validity condition, the accuracy was significantly lower for the invalid cue trials than for the neutral cue trials. Günseli et al. (2015) similarly found a minor detriment in memory quality from invalid retro-cues under the low-validity condition. In our study, the retro-cue costs under the low-validity condition could also be explained by the protection-during-retrieval hypothesis (Makovski & Jiang, 2007, 2008; Makovski & Pertzov, 2015; Makovski, Sussman, & Jiang, 2008), which suggests that the redirection of selective attention to the cued item makes VWM representations more resistant to visual interference from the probe stimulus. Thus, the non-cued items that are not protected by prolonged selective attention are more vulnerable to impairment after the probe stimulus appears, resulting in retro-cue costs from the invalid cues. Notably, we used a change-detection task. The probe stimulus was presented in the same position as the memory item, which led to new visual interference when the probe array appeared (Makovski & Jiang, 2007).

Taken together, our study findings provide new EEG evidence for different mechanisms underlying the retro-cue effect. Importantly, these mechanisms are determined by the expectation of cue validity and by the chosen experimental parameters.

5. Conclusion

Our study results suggest that the maintenance of non-cued representations in VWM is affected by the cue validity. When the retro-cue validity is high, individuals will drop the non-cued items from VWM to strengthen their maintenance of the cued item. By contrast, when the retro-cue validity is low, individuals are more likely to maintain both the cued and non-cued items in VWM, but they prioritize the cued item by attention. The maintenance of the non-cued representations in VWM may be driven in part by strategy or it may be a result of implicit statistical learning. Our study provides new EEG evidence for the previous hypotheses of the retro-cue effect and reconciles previous discrepant results. This research provides an important theoretical basis for further exploration of the relationship between attention and working memory.

Funding

This work was supported by grants from the National Natural Science Foundation of China (no. 31970989 and no. 31700948) to Qiang Liu, the Academy of Finland (no. 333649) to Chaoxiong Ye, and from the program of China Scholarship Council (CSC202008210294) to Xueying Fu. The funders had no role in study design, data collection and analysis, decision to publish or preparation of the manuscript.

Data Availability

The datasets generated/analyzed during this study and experimental script have been added to https://osf.io/qtwc9/.

Author contributions

Xueying Fu: Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation, Resources, Data curation, Visualization, Project administration, Writing - Original Draft, Writing - Review & Editing; **Chaoxiong Ye**: Conceptualization, Methodology, Resources, Validation, Formal analysis, Investigation, Project administration, Funding acquisition, Writing - Original Draft, Writing - Review & Editing; **Zhonghua Hu**: Supervision, Writing - Review & Editing; **Ziyuan Li**: Writing - Review & Editing; **Tengfei Liang**: Formal analysis; **Qiang Liu**: Conceptualization, Supervision, Funding acquisition, Writing - Review & Editing.

References

Allon, A. S., & Luria, R. (2019). Filtering performance in visual working memory is improved by reducing early spatial attention to the distractors. *Psychophysiology, 56*(5), e13323. doi: 10.1111/psyp.13323

Bacigalupo, F., & Luck, S. J. (2019). Lateralized Suppression of Alpha-Band EEG Activity As a Mechanism of Target Processing. *J Neurosci, 39*(5), 900-917. doi: 10.1523/JNEUROSCI.0183-18.2018

Bae, G. Y., & Luck, S. J. (2018). Dissociable Decoding of Spatial Attention and Working Memory from EEG Oscillations and Sustained Potentials. *J Neurosci, 38*(2), 409-422. doi: 10.1523/JNEUROSCI.2860-17.2017

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. doi: 10.1111/j.2517-6161.1995.tb02031.x

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. *Q J Exp Psychol (Hove), 65*(3), 426-438. doi: 10.1080/17470218.2011.604786

Christophel, T. B., Iamshchinina, P., Yan, C., Allefeld, C., & Haynes, J. D. (2018). Cortical specialization for attended versus unattended working memory. *Nat Neurosci, 21*(4), 494-496. doi: 10.1038/s41593-018-0094-4

Christophel, T. B., Klink, P. C., Spitzer, B., Roelfsema, P. R., & Haynes, J. D. (2017). The Distributed Nature of Working Memory. *Trends Cogn Sci, 21*(2), 111-124. doi: 10.1016/j.tics.2016.12.007

Delorme, A., & Makeig, S. (2004). EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *J Neurosci Methods, 134*(1), 9-21. doi: 10.1016/j.jneumeth.2003.10.009

Eimer, M. (1996). The N2pc component as an indicator of attentional selectivity. *Electroencephalogr Clin Neurophysiol, 99*(3), 225-234. doi: 10.1016/0013-4694(96)95711-9

Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G*Power 3: a flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behav Res Methods, 39*(2), 175-191. doi: 10.3758/bf03193146

Feldmann-Wustefeld, T., & Vogel, E. K. (2019). Neural Evidence for the Contribution of Active Suppression During Working Memory Filtering. *Cereb Cortex, 29*(2), 529-543. doi: 10.1093/cercor/bhx336

Feldmann-Wustefeld, T., Vogel, E. K., & Awh, E. (2018). Contralateral Delay Activity Indexes Working Memory Storage, Not the Current Focus of Spatial Attention. *J Cogn Neurosci, 30*(8), 1185-1196. doi: 10.1162/jocn_a_01271

Fukuda, K., Mance, I., & Vogel, E. K. (2015). α power modulation and event-related slow wave provide dissociable correlates of visual working memory. *J Neurosci, 35*(41), 14009-14016. doi: 10.1523/Jneurosci.5003-14.2015

Günseli, E., Fahrenfort, J. J., van Moorselaar, D., Daoultzis, K. C., Meeter, M., & Olivers, C. N. L. (2019). EEG dynamics reveal a dissociation between storage and selective attention within working memory. *Sci Rep, 9*. doi: 10.1038/S41598-019-49577-0

Günseli, E., Meeter, M., & Olivers, C. N. L. (2014). Is a search template an ordinary working memory? Comparing electrophysiological markers of working memory maintenance for visual search and recognition. *Neuropsychologia, 60*, 29-38. doi: j.neuropsychologia.2014.05.012

Günseli, E., Olivers, C. N. L., & Meeter, M. (2014). Effects of search difficulty on the selection, maintenance, and learning of attentional templates. *J Cogn Neurosci, 26*(9), 2042-2054. doi: 10.1162/jocn_a_00600

Günseli, E., van Moorselaar, D., Meeter, M., & Olivers, C. N. (2015). The reliability of retro-cues determines the fate of noncued visual working memory representations. *Psychon Bull Rev*. doi: 10.3758/s13423-014-0796-x

Gao, Z., Xu, X., Chen, Z., Yin, J., Shen, M., & Shui, R. (2011). Contralateral delay activity tracks object identity information in visual short term memory. *Brain Res, 1406*, 30-42. doi: 10.1016/j.brainres.2011.06.049

Gao, Z., Yin, J., Xu, H., Shui, R., & Shen, M. (2011). Tracking object number or information load in visual working memory: revisiting the cognitive implication of contralateral delay activity. *Biol Psychol, 87*(2), 296-302. doi: 10.1016/j.biopsycho.2011.03.013

Goddertz, A., Klatt, L. I., Mertes, C., & Schneider, D. (2018). Retroactive attentional shifts predict performance in a working memory task: Evidence by lateralized EEG patterns. *Front Hum Neurosci, 12*, 428. doi: 10.3389/fnhum.2018.00428

Griffin, I. C., & Nobre, A. C. (2003). Orienting attention to locations in internal representations. *J Cogn Neurosci, 15*(8), 1176-1194. doi: 10.1162/089892903322598139

Groppe, D. M., Urbach, T. P., & Kutas, M. (2011). Mass univariate analysis of event-related brain potentials/fields I: A critical tutorial review. *Psychophysiology, 48*(12), 1711-1725. doi: j.1469-8986.2011.01273.x

Hakim, N., Adam, K. C. S., Günseli, E., Awh, E., & Vogel, E. K. (2019). Dissecting the Neural Focus of Attention Reveals Distinct Processes for Spatial Attention and Object-Based Storage in Visual Working Memory. *Psychol Sci, 30*(4), 526-540. doi: 10.1177/0956797619830384

Hakim, N., Feldmann-Wüstefeld, T., Awh, E., & Vogel, E. K. (2020). Perturbing neural representations of working memory with task-irrelevant interruption. *J Cogn Neurosci, 32*(3), 558-569. doi: 10.1162/jocn_a_01481

Hao, R., Becker, M. W., Ye, C., Liu, Q., & Liu, T. (2018). The bandwidth of VWM consolidation varies with the stimulus feature: Evidence from event-related potentials. *J Exp Psychol Hum Percept Perform, 44*(5), 767-777. doi: 10.1037/xhp0000488

Ikkai, A., Dandekar, S., & Curtis, C. E. (2016). Lateralization in Alpha-Band Oscillations Predicts the Locus and Spatial Distribution of Attention. *PLoS One, 11*(5), e0154796. doi: 10.1371/journal.pone.0154796

Ikkai, A., McCollough, A. W., & Vogel, E. K. (2010). Contralateral delay activity provides a neural measure of the number of representations in visual working memory. *J Neurophysiol, 103*(4), 1963-1968. doi: 10.1152/jn.00978.2009

Keshvari, S., van den Berg, R., & Ma, W. J. (2013). No Evidence for an Item Limit in Change Detection. *PLOS Computational Biology, 9*(2), e1002927.

Kiss, M., Van Velzen, J., & Eimer, M. (2008). The N2pc component and its links to attention shifts and spatially selective visual processing. *Psychophysiology, 45*(2), 240-249. doi: 10.1111/j.1469-8986.2007.00611.x

Klatt, L. I., Getzmann, S., Wascher, E., & Schneider, D. (2018). Searching for auditory targets in external space and in working memory: Electrophysiological mechanisms underlying perceptual and retroactive spatial attention. *Behav Brain Res, 353*, 98-107. doi: 10.1016/j.bbr.2018.06.022

Kuo, B. C., Stokes, M. G., & Nobre, A. C. (2012). Attention modulates maintenance of representations in visual short-term memory. *J Cogn Neurosci, 24*(1), 51-60. doi: 10.1162/jocn_a_00087

Landman, R., Spekreijse, H., & Lamme, V. A. (2003). Large capacity storage of integrated objects before change blindness. *Vision Res, 43*(2), 149-164. doi: 10.1016/S0042-6989(02)00402-9

Liang, T., Cheng, Z., Hu, W., Ye, C., Zhang, J., & Liu, Q. (2020). Limitations of concurrently representing objects within view and in visual working memory. *Sci Rep, 10*, 5351. doi: 10.1038/s41598-020-62164-y

Luck, S. J., & Hillyard, S. A. (1994a). Electrophysiological correlates of feature analysis during visual search. *Psychophysiology, 31*(3), 291-308. doi: 10.1111/j.1469-8986.1994.tb02218.x

Luck, S. J., & Hillyard, S. A. (1994b). Spatial filtering during visual search: evidence from human electrophysiology. *J Exp Psychol Hum Percept Perform, 20*(5), 1000-1014. doi: 10.1037//0096-1523.20.5.1000

Luria, R., Balaban, H., Awh, E., & Vogel, E. K. (2016). The contralateral delay activity as a neural measure of visual working memory. *Neurosci Biobehav Rev, 62*, 100-108. doi: 10.1016/j.neubiorev.2016.01.003

Ma, W. J., Husain, M., & Bays, P. M. (2014). Changing concepts of working memory. *Nat Neurosci, 17*(3), 347-356. doi: 10.1038/nn.3655

Makovski, T., & Jiang, Y. V. (2007). Distributing versus focusing attention in visual short-term memory. *Psychon Bull Rev, 14*(6), 1072-1078.

Makovski, T., & Jiang, Y. V. (2008). Proactive interference from items previously stored in visual working memory. *Mem Cognit, 36*(1), 43-52.

Makovski, T., & Pertzov, Y. (2015). Attention and memory protection: Interactions between retrospective attention cueing and interference. *Q J Exp Psychol (Hove), 68*(9), 1735-1743. doi: 10.1080/17470218.2015.1049623

Makovski, T., Sussman, R., & Jiang, Y. V. (2008). Orienting attention in visual working memory reduces interference from memory probes. *J Exp Psychol Learn Mem Cogn, 34*(2), 369-380. doi: 10.1037/0278-7393.34.2.369

McCollough, A. W., Machizawa, M. G., & Vogel, E. K. (2007). Electrophysiological measures of maintaining representations in visual working memory. *Cortex, 43*(1), 77-94. doi: 10.1016/S0010-9452(08)70447-7

Myers, N. E., Chekroud, S. R., Stokes, M. G., & Nobre, A. C. (2018). Benefits of flexible prioritization in working memory can arise without costs. *J Exp Psychol Hum Percept Perform, 44*(3), 398. doi: 10.1037/xhp0000449

Ngiam, W. X. Q., Adam, K. C. S., Quirk, C., Vogel, E. K., & Awh, E. (2021). Estimating the

statistical power to detect set-size effects in contralateral delay activity. *Psychophysiology, 58*(5), e13791. doi: 10.1111/psyp.13791
Peterson, D. J., Gozenman, F., Arciniega, H., & Berryhill, M. E. (2015). Contralateral delay activity tracks the influence of Gestalt grouping principles on active visual working memory representations. *Atten Percept Psychophys*. doi: 10.3758/s13414-015-0929-y
Poch, C., Capilla, A., Hinojosa, J. A., & Campo, P. (2017). Selection within working memory based on a color retro-cue modulates alpha oscillations. *Neuropsychologia, 106*, 133-137. doi: 10.1016/j.neuropsychologia.2017.09.027
Poch, C., Valdivia, M., Capilla, A., Hinojosa, J. A., & Campo, P. (2018). Suppression of no-longer relevant information in Working Memory: An alpha-power related mechanism? *Biol Psychol, 135*, 112-116. doi: 10.1016/j.biopsycho.2018.03.009
Rouder, J. N., Speckman, P. L., Sun, D., Morey, R. D., & Iverson, G. (2009). Bayesian t tests for accepting and rejecting the null hypothesis. *Psychon Bull Rev, 16*(2), 225-237. doi: 10.3758/PBR.16.2.225
Salahub, C., & Emrich, S. M. (2020). Fear not! Anxiety biases attentional enhancement of threat without impairing working memory filtering. *Cogn Affect Behav Neurosci, 20*(6), 1248-1260. doi: 10.3758/s13415-020-00831-3
Sauseng, P., Klimesch, W., Stadler, W., Schabus, M., Doppelmayr, M., Hanslmayr, S., . . . Birbaumer, N. (2005). A shift of visual spatial attention is selectively associated with human EEG alpha activity. *Eur J Neurosci, 22*(11), 2917-2926. doi: 10.1111/j.1460-9568.2005.04482.x
Schmalz, X., Biurrun Manresa, J., & Zhang, L. (2021). What is a Bayes factor? *Psychological Methods*. doi: 10.1037/met0000421
Schneider, D., Barth, A., Getzmann, S., & Wascher, E. (2017). On the neural mechanisms underlying the protective function of retroactive cuing against perceptual interference: Evidence by event-related potentials of the EEG. *Biol Psychol, 124*, 47-56. doi: 10.1016/j.biopsycho.2017.01.006
Souza, A. S., & Oberauer, K. (2016). In search of the focus of attention in working memory: 13 years of the retro-cue effect. *Atten Percept Psychophys, 78*(7), 1839-1860. doi: 10.3758/s13414-016-1108-5
Stevanovski, B., & Jolicoeur, P. (2011). Consolidation of multifeature items in visual working memory: central capacity requirements for visual consolidation. *Atten Percept Psychophys, 73*(4), 1108-1119. doi: 10.3758/s13414-011-0099-5
Thut, G., Nietzel, A., Brandt, S. A., & Pascual-Leone, A. (2006). Alpha-band electroencephalographic activity over occipital cortex indexes visuospatial attention bias and predicts visual target detection. *J Neurosci, 26*(37), 9494-9502. doi: 10.1523/JNEUROSCI.0875-06.2006
van Dijk, H., van der Werf, J., Mazaheri, A., Medendorp, W. P., & Jensen, O. (2010). Modulations in oscillatory activity with amplitude asymmetry can produce cognitively relevant event-related responses. *Proc Natl Acad Sci U S A, 107*(2), 900-905. doi: 10.1073/pnas.0908821107
Vogel, E. K., & Machizawa, M. G. (2004). Neural activity predicts individual differences in visual working memory capacity. *Nature, 428*(6984), 748-751. doi: 10.1038/nature02447

Vogel, E. K., McCollough, A. W., & Machizawa, M. G. (2005). Neural measures reveal individual differences in controlling access to working memory. *Nature, 438*(7067), 500-503. doi: 10.1038/nature04171

Vogel, E. K., Woodman, G. F., & Luck, S. J. (2006). The time course of consolidation in visual working memory. *J Exp Psychol Hum Percept Perform, 32*(6), 1436-1451. doi: 10.1037/0096-1523.32.6.1436

Wang, S., Rajsic, J., & Woodman, G. F. (2019). The Contralateral Delay Activity Tracks the Sequential Loading of Objects into Visual Working Memory, Unlike Lateralized Alpha Oscillations. *J Cogn Neurosci, 31*(11), 1689-1698. doi: 10.1162/jocn_a_01446

Williams, M., Hong, S. W., Kang, M. S., Carlisle, N. B., & Woodman, G. F. (2013). The benefit of forgetting. *Psychon Bull Rev, 20*(2), 348-355. doi: 10.3758/s13423-012-0354-3

Williams, M., & Woodman, G. F. (2012). Directed forgetting and directed remembering in visual working memory. *J Exp Psychol Learn Mem Cogn, 38*(5), 1206-1220. doi: 10.1037/a0027389

Wolff, M. J., Jochim, J., Akyurek, E. G., & Stokes, M. G. (2017). Dynamic hidden states underlying working-memory-guided behavior. *Nat Neurosci, 20*(6), 864-871. doi: 10.1038/nn.4546

Worden, M. S., Foxe, J. J., Wang, N., & Simpson, G. V. (2000). Anticipatory biasing of visuospatial attention indexed by retinotopically specific alpha-band electroencephalography increases over occipital cortex. *J Neurosci, 20*(6), RC63.

Ye, C., Hu, Z., Li, H., Ristaniemi, T., Liu, Q., & Liu, T. (2017). A two-phase model of resource allocation in visual working memory. *J Exp Psychol Learn Mem Cogn, 43*(10), 1557-1566. doi: 10.1037/xlm0000376

Ye, C., Liang, T., Zhang, Y., Xu, Q., Zhu, Y., & Liu, Q. (2020). The two-stage process in visual working memory consolidation. *Sci Rep, 10*(1), 13564. doi: 10.1038/s41598-020-70418-y

Ye, C., Sun, H. J., Xu, Q., Liang, T., Zhang, Y., & Liu, Q. (2019). Working memory capacity affects trade-off between quality and quantity only when stimulus exposure duration is sufficient: Evidence for the two-phase model. *Sci Rep, 9*(1), 8727. doi: 10.1038/s41598-019-44998-3

Ye, C., Xu, Q., Liu, Q., Cong, F., Saariluoma, P., Ristaniemi, T., & Astikainen, P. (2018). The impact of visual working memory capacity on the filtering efficiency of emotional face distractors. *Biol Psychol, 138*, 63-72. doi: 10.1016/j.biopsycho.2018.08.009

Ye, C., Zhang, L., Liu, T., Li, H., & Liu, Q. (2014). Visual working memory capacity for color is independent of representation resolution. *PLoS One, 9*(3), e91681. doi: 10.1371/journal.pone.0091681

Zhang, L., Peng, W., Zhang, Z., & Hu, L. (2013). Distinct features of auditory steady-state responses as compared to transient event-related potentials. *PLoS One, 8*(7), e69164. doi: 10.1371/journal.pone.0069164

Zhang, Z. G., Hu, L., Hung, Y. S., Mouraux, A., & Iannetti, G. D. (2012). Gamma-band oscillations in the primary somatosensory cortex--a direct and obligatory correlate of subjective pain intensity. *J Neurosci, 32*(22), 7429-7438. doi: 10.1523/JNEUROSCI.5877-11.2012

Zhao, G., Liu, Q., Zhang, Y., Jiao, J., Zhang, Q., Sun, H., & Li, H. (2011). The amplitude of N2pc reflects the physical disparity between target item and distracters. *Neurosci Lett, 491*(1), 68-72. doi: 10.1016/j.neulet.2010.12.066

Zivony, A., Allon, A. S., Luria, R., & Lamy, D. (2018). Dissociating between the N2pc and attentional shifting: An attentional blink study. *Neuropsychologia, 121*, 153-163. doi: 10.1016/j.neuropsychologia.2018.11.003