Is There an Activity-silent Working Memory?

Oberauer, Klaus ; Awh, Edward

DOI: https://doi.org/10.1162/jocn_a_01917

Originally published at:
Oberauer, Klaus; Awh, Edward (2022). Is There an Activity-silent Working Memory? Journal of Cognitive Neuroscience, 34(12):2360-2374.
DOI: https://doi.org/10.1162/jocn_a_01917

Is there an activity-silent working memory?

Klaus Oberauer, University of Zurich

Edward Awh, University of Chicago

The research reported in this article was supported by a grant from the Swiss National Science Foundation (grant number 100014_126766) to Klaus Oberauer. All raw data and analysis code can be found on the OSF: https://osf.io/tp49k/

Correspondence should be addressed to: Klaus Oberauer, University of Zurich, Department of Psychology – Cognitive Psychology; Binzmühlestrasse 14/22, 8050 Zürich, Switzerland, Email: k.oberauer@psychologie.uzh.ch

Abstract

Although storage in working memory (WM) can be tracked via measurements of ongoing neural activity, past work has shown that observers can maintain access to that information despite temporary interruptions of those neural patterns. This observation has been regarded as evidence for a neurally silent form of WM storage. Alternatively, however, unattended information could be retrieved from episodic long-term memory (eLTM) rather than being maintained in WM during the activity-silent period. Here, we tested between these possibilities by examining whether WM performance showed evidence of proactive interference – a hallmark of retrieval from eLTM -- following such interruptions. Participants remembered the colors (Exp. 1-3) or locations (Exp. 4) of serially presented objects. We found PI for set sizes larger than 4, but not for smaller set sizes, suggesting that eLTM may have supported performance when WM capacity was exceeded. Critically, performance with small set sizes remained resistant to PI, even following prolonged interruptions by a challenging distractor task. Thus, we found evidence for PI-resistant memories that were maintained across likely interruptions of storage-related neural activity, an empirical pattern that implies activity-silent storage in WM.

Is there an activity-silent working memory?

Working memory refers to our ability to hold a small amount of information available for ongoing cognitive processes. Two features distinguish working memory (WM) from episodic long-term memory (eLTM): First, WM has a severely limited capacity – typically no more than three or four distinct representational units, or chunks (Cowan, 2001). Second, the contents of WM are protected against proactive interference from information in eLTM (Oberauer, Awh, & Sutterer, 2017; Oberauer & Greve, 2021).

Advances in neuroscience have made it possible to decode the contents of WM from the spatial distribution of BOLD or EEG signals during the maintenance interval of a WM task (Christophel, Hebart, & Haynes, 2012; Foster, Sutterer, Serences, Vogel, & Awh, 2016; Harrison & Tong, 2009; Serences, Ester, Vogel, & Awh, 2009). Using these techniques, it has been shown that when observers are cued to selectively attend to a subset of the stored items, that the decoding of unattended items often drops to baseline (Hakim, Feldmann-Wüstefeld, Awh, & Vogel, 2021; LaRocque, Lewis-Peacock, Drysdale, Oberauer, & Postle, 2013; Rose et al., 2016; van Moorselaar et al., 2017). In addition, when attention is directed back to the unattended items, the active neural traces of those items are often restored (LaRocque et al., 2013; Lewis-Peacock, Drysdale, Oberauer, & Postle, 2011; Sprague, Ester, & Serences, 2016; van Moorselaar et al., 2017). Finally, unattended information in WM can be made decodable by uninformative visual or neural impulses (Rose et al., 2016; Wolff, Ding, Myers, & Stokes, 2015), suggesting a lingering pattern of synaptic connectivity that may enable maintenance of working memories in an “activity-silent” manner (Stokes, 2015).

The idea that WM representations are maintained through short-term changes to synaptic weights is in agreement with a number of computational models of WM, in which memory sets are maintained by binding each item to its context, and context cues are used to re-activate item representations for retrieval (Botvinick & Watanabe, 2007; Burgess & Hitch, 1999; Lewandowsky & Farrell, 2008; Oberauer, Lewandowsky, Farrell, Jarrold, & Greaves, 2012; Oberauer & Lin, 2017).

However, an alternative interpretation of neurally silent memory states is that currently unattended information is not held in WM at all, but maintained only in eLTM (Beukers, Buschman, Cohen, & Norman, 2021; Foster, Vogel, & Awh, in press). Although it is clear that such episodic memories exist, it is less clear what role episodic LTM may play in demonstrations of activity-silent WM.

The aim of the present work is to adjudicate between these two hypotheses. Our rationale is that information in an activity-silent WM should be protected against proactive interference (PI) from preceding events (Oberauer et al., 2017), whereas information in eLTM should not. Thus, we considered resistance to PI to be the “fingerprints” of storage in WM. Our experiments build on previous work showing that a simple distractor task (van Moorselaar et al., 2017), and even the presentation of distracting stimuli without any associated task (Hakim, Feldmann-Wüstefeld, Awh, & Vogel, 2019), abolishes some well-established neural signatures of WM maintenance, but only modestly impairs memory performance. Thus, such tasks provide a testbed for examining how memory performance is affected by the interruption of storage-related neural activity.

We tested WM with and without distraction during the retention interval, and varied the level of PI between trials. In addition, we varied the size of the memory set, with smaller sets within WM capacity (about 3 items) and larger sets clearly exceeding it so that contributions from eLTM were expected. We predicted that without distraction, retrieval of small sets should be unaffected by PI, because it can rely on WM, whereas retrieval of larger sets should show PI, because of contributions from eLTM. The key question is: Does an intervening task known to interrupt active neural signatures of WM storage force participants to rely on eLTM even at small set sizes?

If that is the case, then performance in the distraction condition should become vulnerable to PI even with a small set size. By contrast, if the distractor task merely draws attention away from WM contents, which are meanwhile held in an activity-silent state in WM, then performance on small memory sets should still be unaffected by PI.

Experiment 1

With Experiment 1 we test the prediction that, to the extent that the memory set size exceeds WM capacity, performance becomes more reliant on eLTM. If that is the case, then performance in a WM test should become more vulnerable to PI when the set size exceeds WM capacity.

We used the task introduced by Brady, Konkle, Gill, Oliva, and Alvarez (2013) for testing memory for arbitrary object-color relations (see Figure 1), because it is suitable for manipulating PI between trials. Participants were presented with 1 to 8 pictures of concrete objects uniformly colored in a color drawn at random from a color circle. Serial presentations ensured that all memoranda were focally attended and encoded without competition from other items; thus, even if capacity limits prevented the concurrent maintenance of all items in working memory, eLTM traces of all items should have been encoded. At test, participants had to reproduce the color of each presented object by choosing it on the color circle. In the condition with high PI, a small set of objects was used repeatedly between trials, always with a new random color. In the condition with low PI, each object was used only once throughout the experiment.

Method

Participants. Sixty students from the University of Zurich, between 18 and 35 years old, took part in a one-hour session in exchange for partial course credit or 15 Swiss Francs (about 15 US dollars). They were allocated at random to four groups created by crossing PI (high vs. low) with presentation time (short vs. long), resulting in N = 15 per group. The sample size and number of trials for this and the subsequent experiments were initially set as a compromise between available resources and our estimate of what was required to measure medium-sized effects with this experimental paradigm. As we planned to use Bayesian analysis methods, we anticipated that we could increase the sample size in case tests of theoretically important hypotheses were ambiguous. This turned out to be unnecessary. The experiments were carried out in accordance with the ethical guidelines of the Faculty of Arts and Social Sciences at the University of Zurich. All participants signed a consent form after being informed about the experiment, and they were debriefed about the purpose of the experiment after completion.

Design. PI (high vs. low) and presentation time (300 vs. 1000 ms inter-item interval) were varied independently between subjects. Set size and delay were varied within subjects. Memory sets consisted of 1, 2, 3, 4, 6, or 8 objects. We varied the delay between presentation and test for the smaller set sizes because with sequential presentation, the average retention interval between presentation and test increases with set size. For set sizes 1 and 2, we therefore used three delays. The shortest was set to the standard delay of 1 s between offset of the last object and onset of the first test. The longest was set to 8 times the inter-stimulus interval (ISI) plus 1 s (set size 1), or 7 x ISI + 1 s (set size 2) to approximately equate the average retention interval of the small set sizes to the average retention interval of the largest set size of eight objects. The intermediate delay was 4 x ISI + 1 s. For set size 3 there were two delays, one of 1 s, and the other set to 6 x ISI + 1 s. The higher set sizes were run only with the standard 1 s delay.

Materials and Procedure. We used the 385 silhouette images of concrete nameable objects of Oberauer et al. (2017) and combined them with colors drawn at random from 360 colors on a color circle in the CIE L x a x b color space, centered on L = 70, a = 20, and b = 38, with a radius of 60. The silhouette images were presented uniformly in the chosen color against a grey background. For the condition with high PI, a random set of eight objects was selected as the object pool for each participant, and the objects for each trial were a random subset from that pool. For the condition with low PI, the objects for each trial were sampled without replacement from the pool of all 385 objects, until all objects were used once; then the full pool was reinstated for a second round of sampling. For each trial, the objects were given new random colors, and presented in a random order. In this way, in the high proactive-interference condition the same objects appeared repeatedly across trials, always with a new color, creating a condition known to produce strong PI. In the low proactive-interference condition, objects re-appeared very rarely, and only after many trials with other objects, creating a condition with low PI between trials.

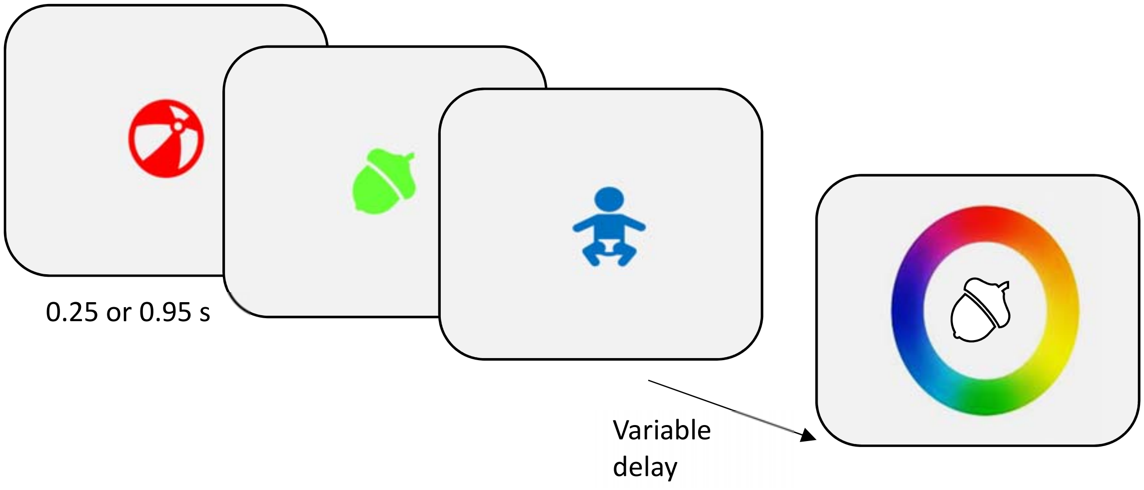

Figure 1: Illustration of a trial with set size 3. Objects with random colors were presented sequentially. After a variable delay, all objects were tested in a random order (only the first test is shown). The tested object was presented in white, and participants selected the color they remembered from the color wheel.

Each trial began with a fixation dot presented for 50 ms in the screen center, followed by the sequential presentation of the colored objects. The objects were presented for 250 ms (fast) or 950 ms (slow) in the screen center, with a blank inter-stimulus interval of 50 ms. After offset of the last object, and the delay, the test phase began with the presentation of the fixation dot for 250 ms. Then one of the objects was presented in white in the screen center, surrounded by a color wheel. Participants were asked to select the color they remembered for that object by clicking on the appropriate color in the wheel, upon which the screen went blank apart from the fixation dot. All objects were tested in this way in random order.[1] For each tested object the color wheel was

---

[1] In experiments on visual WM, researchers often test only one array item, whereas in experiments on verbal WM, all items are usually tested. We tested all items to obtain more information per trial.

presented in a new random orientation. After the final test, participants started the next trial by pressing the space bar. Figure 1 shows an example trial.

The experiment was organized into two blocks. Each block started with 5 practice trials (one for each set-size condition, with the intermediate delays for set sizes 1 and 2, and the standard delay for all others) followed by 110 test trials – 10 for each of the combinations of set size with delay. The practice trials were given in order of ascending set size; the test trials were presented in a random order.

Data Analysis

As dependent variable we computed the absolute error of color reproduction in degrees. We analyzed the data with Bayesian linear mixed models run with the *brms* package (Bürkner, 2017). For analysis, the dependent variable and all independent variables were z-standardized so that all estimated effects were on a standard effect-size scale. This enabled us to use the default priors for standardized effect sizes recommended by Rouder, Morey, Speckman, and Province (2012), except that we standardized the effect sizes through division by the standard deviations of the variables rather than by the standard deviation of the residual (see Singmann et al., 2021, for why standardization by the residual is problematic).

Models were run with a Cauchy prior with scale = 0.5 on fixed effects. Predictors varying parametrically over more than two levels (i.e., set size, delay) were modelled as linear effects; predictors with two nominal levels were coded as effect contrasts (-1 vs. 1). All models were applied to unaggregated data. We tested the effects of interest through model comparison through Bayes factors (BFs), which we estimated with the bridge sampler (Gronau, Singmann, & Wagenmakers, 2018). We started with the full model including all fixed effects, as well as random effects of subject on the intercept and on the within-subjects effects (i.e., random slopes) of the model (Oberauer, 2022). In a first step we tested the random slopes by comparing the full model to a model in which some or all random slopes were removed; for the subsequent steps we maintained the random slopes that received support through the BF.

We next tested fixed effects one by one by comparing the current model to a constrained model obtained by removing the fixed effect of interest from it but keeping the corresponding effect in the random-effect structure (unless it had been removed before). The BF in favor of the unconstrained model reflects the strength of evidence in favor of the tested effect. A $BF_{10} > 1$ is evidence for the effect, whereas a $BF_{10} < 1$ is evidence against the effect, and the BF in favor of the null, BF_{01}, can be obtained as its reciprocal: $BF_{01} = 1/BF_{10}$. We started with testing the highest-order interaction, then progressing to lower-order interactions and finally main effects. At each step we removed effects from the model when the evidence was against them but maintained them when the evidence was in their favor; the resulting model served as the starting point for subsequent testing of lower-order effects. Table 1 summarizes the BFs from all linear mixed models for Experiment 1 up to the 2-way interactions (3-way interactions are reported in the text).

Results

Figure 2 shows the mean errors as a function of set size and PI, split by the two presentation times. It is clear that with the short presentation time there was no effect of PI. With the longer presentation time, PI impaired memory at set sizes larger than 4. Analysis 1 applied a model to the full data including set size, PI, and presentation time as predictors. It confirmed the obvious main effects of set size and presentation time, and provided moderate evidence for the predicted interaction of PI and set size. Although the evidence pertaining to the 3-way interaction had an ambiguous BF_{10} of 1.08, it is clear from Figure 2 that there was PI only in the slow-presentation group. Therefore, Analysis 2 zoomed in on that subset of the data. In the group with longer presentation time, there was clear evidence for the predicted interaction of set size and PI.

As there was no evidence for a main effect of PI across set sizes, we focused on the two set sizes clearly exceeding WM capacity (i.e., 6 and 8 objects) for the longer presentation time. For that subset of conditions a Bayesian t-test (Morey & Rouder, 2015) showed clear evidence for PI, BF_{10} = 43.4.

Table 1: Bayes Factors (BF_{10}) from tests of fixed effects in Experiment 1

Analysis 1	**Main Effect**	**Interaction with PI**	**Interaction with Presentation Time**
Set size	7.3×10^{38}	7.0	790
PI	0.045	--	0.1
Presentation time	455	--	--
Analysis 2 (Long presentation time)	**Main Effect**	**Interaction with PI**	
Set size	4.4×10^{16}	41	
PI	0.14		
Analysis 3 (Long presentation time, set sizes 1-3)	**Main Effect**	**Interaction with PI**	**Interaction with Delay**
Set size	3.2×10^{131}	0.06	1.05
PI	0.06		0.32
Delay	0.44		

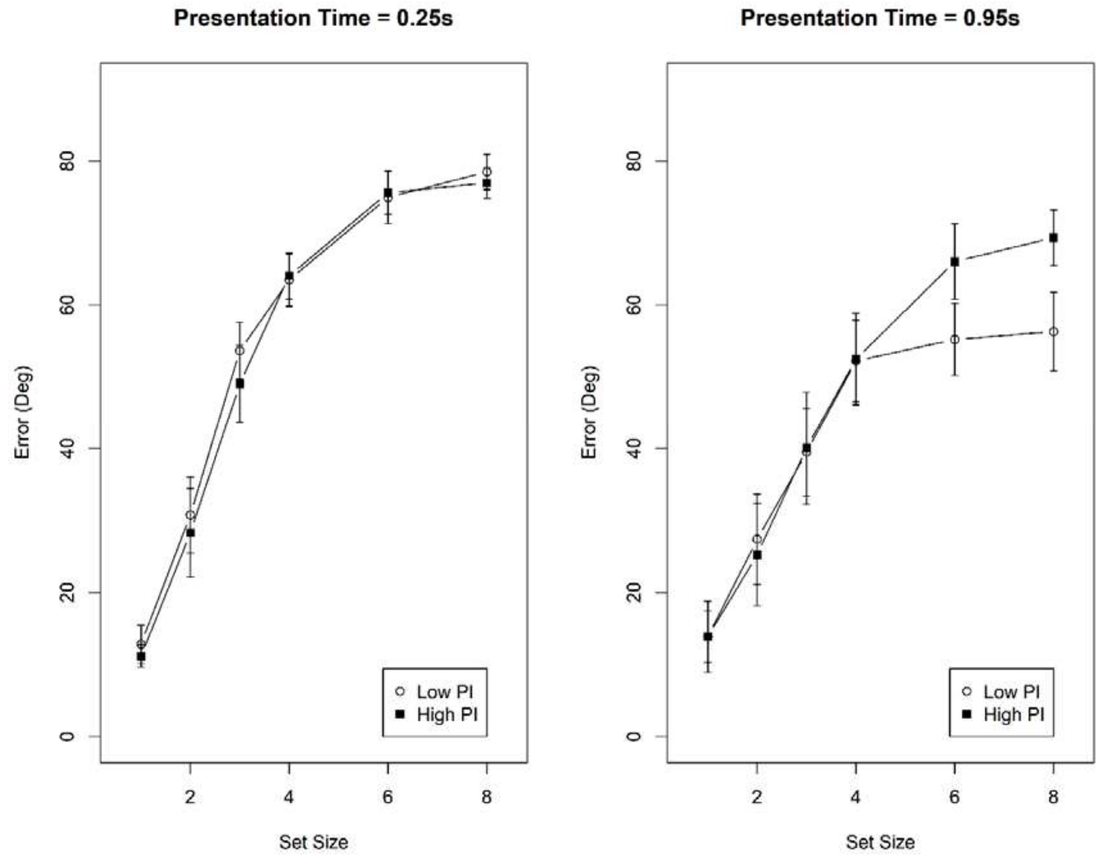

Figure 2: Experiment 1: Mean error of reproduction as a function of set size and PI, for short (left panel) and long (right panel) presentation times. Error bars are 95% confidence intervals.

The finding of PI only at the larger set sizes could be a side effect of the longer average retention interval. This would be in line with temporal-distinctiveness models of memory (Brown, Neath, & Chater, 2007), according to which events are represented in memory on a psychological time dimension that is increasingly compressed further from the present. As a consequence, events with the same objective temporal separation become less distinctive on the time dimension in memory as they recede into the past. In the present experiments, a longer retention interval would make the stimuli of the present trial temporally less distinctive from stimuli of preceding trials. This could lead to more PI between trials. To test this possibility, Analysis 3 focuses on set sizes 1 to 3, for which we varied the retention interval, again using only data from the slow-presentation group. Figure 3 shows the results. Even with delays up to 8 s there was no hint of PI at these set sizes (the reciprocals of the respective BF_{10} in Table 1 are: BF_{01} = 14.7 against the main effect of PI, and BF_{01} = 3.2 against its interaction with delay), contrary to the assumption that the PI we observed for larger set sizes arose from their longer retention interval. There was also evidence against the three-way interaction, BF_{10} = 0.06.

A further way to gauge the effect of PI is through intrusions from previous trials. To the extent that episodic LTM contributes to people's responses, participants should sometimes erroneously retrieve the color that the tested object had on a previous trial. This was indeed the case in the high-PI condition at set sizes 6 and 8 (see the Supplementary document on the OSF project: https://osf.io/tp49k/).

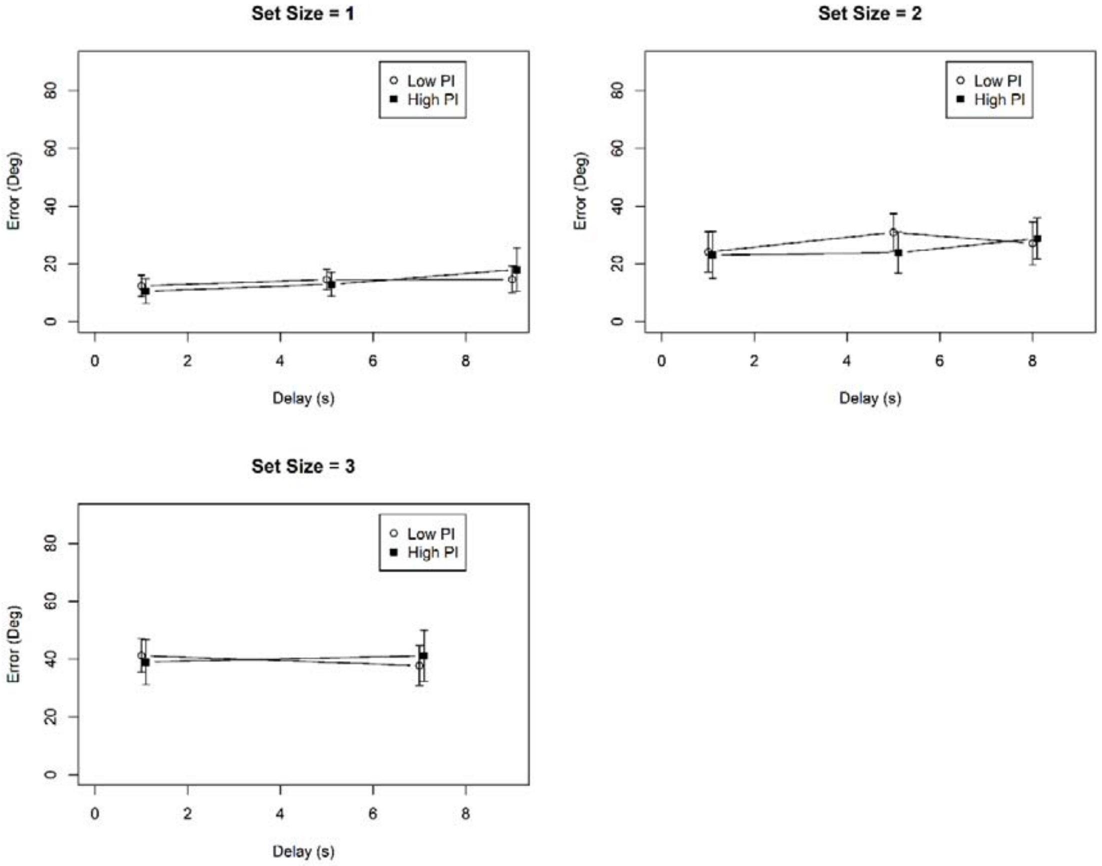

Figure 3: Experiment 1: Mean error of reproduction as a function of delay and PI for the three smallest set sizes, condition with slow presentation time. Error bars are 95% confidence intervals.

Discussion

Experiment 1 confirmed that there is PI between trials if, and only if, the memory set size exceeds WM capacity. We found this only for the longer presentation time. This could be because the short presentation time was not suited for forming sufficiently strong representations in eLTM to be of any use in the task, and as a consequence, people did not try to draw on eLTM, or if they did, they did not retrieve information strong enough to improve performance regardless of the PI manipulation. For the following experiments we therefore used the longer presentation time to investigate whether, after a distractor task that is known to erase a neural signature of maintenance, PI will be found even for smaller set sizes.

Experiments 2 and 3

Experiments 2 and 3 focus on what is probably the empirically best-established neural signature of the load on working memory, the contralateral delay activity (CDA). The CDA is a slow negative deflection of the EEG contralateral to the to-be-remembered memory array. Its amplitude increases with the number of to-be-remembered visual objects up to the person’s WM capacity (Luria, Balaban, Awh, & Vogel, 2016; Vogel, McCollough, & Machizawa, 2005). The CDA is vulnerable to disruption. An array of objects presented during the delay, even when irrelevant to the task, largely abolishes the CDA, but has only a mild effect on memory performance (Hakim et al., 2019). This finding raises the question whether memory in those trials in which the CDA was abolished relied on eLTM, or on a form of “activity-silent” WM that does not evoke the CDA.

We used a subitizing task as the distractor task in the retention interval to disrupt the CDA. Subitizing – the enumeration of small sets of objects – has been shown to generate a CDA of its own, with an asymptote that increases with the set size of the to-be-enumerated set up to four, and then levels off (Mazza & Caramazza, 2015). Therefore, we can be confident that subitizing engages the process that causes the CDA. If that process is the neurally active maintenance of objects in WM, then subitizing 3 to 5 objects should abolish neurally active maintenance of the object-color conjunctions, for two reasons. First, the process reflected by the CDA has a capacity limit of about 3-4 objects, as shown by the fact that the CDA asymptote plateaus with about 3 (for WM maintenance) or 4 (for subitizing) objects, and therefore, the subitizing array should replace the memory array from neurally active maintenance. Second, if even an array of visual objects that people try to ignore disrupts the CDA (Hakim et al., 2019), then an array that people need to attend to should do so, too.

Based on these assumptions, as well as a functional definition of WM that includes PI-resistance, we tested the following predictions: If memory for object-color conjunctions in the absence of the CDA relies on eLTM, then we should observe PI even with small set sizes in the condition with the subitizing distractor task. If, however, small sets of objects are maintained during the interruption via a form of activity-silent WM that does not elicit a CDA, then we should observe PI only for large set sizes, regardless of the distractor condition.

Method

Participants. Thirty-nine (Experiment 2) and 60 (Experiment 3) students from the University of Zurich took part in a one-hour session in exchange for partial course credit or 15 Swiss Francs. They were randomly allocated to two groups.

Design. PI was varied between groups. Set size (3 vs. 8 objects) and distraction condition (without vs. with distractor task) were varied within subjects. The distraction condition was varied between blocks, and set size was varied randomly within blocks. In Experiment 2, a single subitizing trial was to be carried out in the retention interval in the distraction condition. In Experiment 3, we additionally varied the number of subitizing trials in that condition (1, 2, or 4 trials).

Materials and Procedure. The materials and procedure for the memory task were the same as in Experiment 1. In the distraction condition, the (first) subitizing trial started 200 ms after offset of the last memory stimulus with the presentation of a 3 x 3 grid in the upper half of the screen, centered horizontally, and of three response boxes underneath it. The response boxes were the digits 3, 4, and 5, ordered from left to right, each surrounded by a rectangular frame (see Figure 4). Another 200 ms later the nine grid cells were filled with nine black figures – a random mixture of 3, 4, or 5 squares in a random subset of cells, and circles in the remaining cells. Participants' task was to determine as quickly as possible the number of squares, ignoring the circles, and to click on the corresponding response box with the mouse. Once a response option was chosen, the screen went blank.

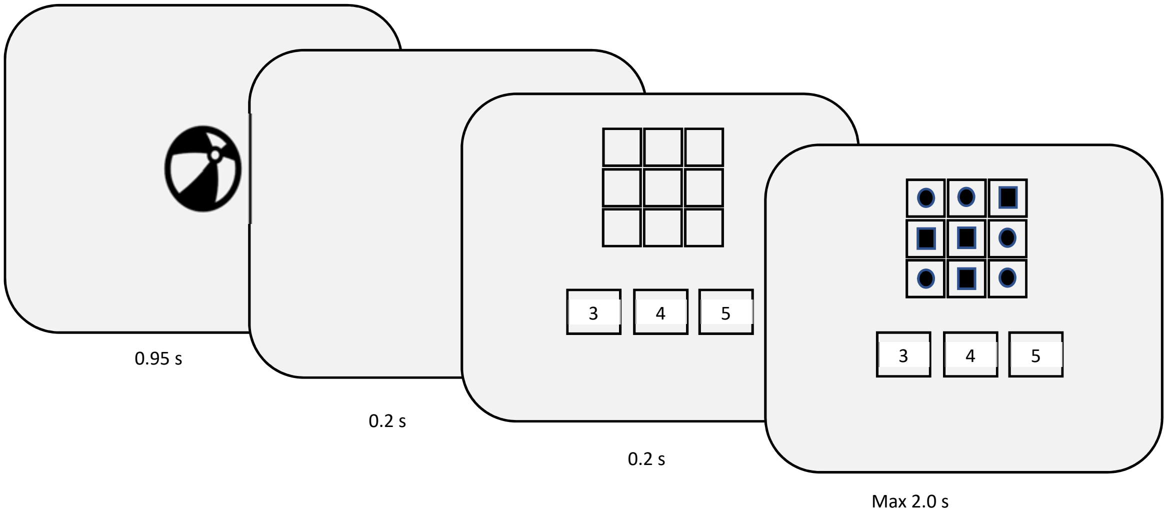

Figure 4: Illustration of a subitizing trial following the last-presented memory stimulus in the distractor conditions in Experiments 2 and 3.

One subitizing trial always lasted 3 s. Participants had a deadline for responding of 2 s after the onset of the subitizing array; when they responded before the deadline, the screen remained blank for the remainder of the time until 3 s have elapsed for the entire trial (starting with the offset of the last memory stimulus); if they did not respond before the deadline, the screen went blank for the remaining time up to 3 s, and the trial was counted as error. In Experiment 3, the first subitizing trial could be followed by one or three more, each again lasting 3 s. To match the retention interval between the distraction and the no-distraction condition, the no-distraction condition had an unfilled delay of 3 s (Experiment 2), or of 3, 6, or 12 s, corresponding to the variation of the number of subitizing trials (Experiment 3).

At the beginning of the experiments, participants practiced the subitizing task on its own for 30 trials. After that they worked through 6 blocks of memory trials, alternating between the distraction and the no-distraction condition; the order of conditions (starting with vs. without distraction in the first block) was counterbalanced across participants. In Experiment 2, each block consisted of 16 test trials (8 per set size). In Experiment 3, each block consisted of 12 test trials, evenly split among trials with 1, 2, and 4 distractor trials, or the corresponding unfilled delay durations. The first two blocks were preceded by 4 practice trials (two per set size).

Results

Figures 5 and 6 show the reproduction errors as a function of set size and PI, separately for the two distraction conditions, for Experiments 2 and 3, respectively. In both experiments, there was PI at the larger but not the smaller set size, regardless of distraction. Tables 2 and 3 summarize the BFs in favor of the fixed effects up to the 2-level interactions for Experiments 2 and 3, respectively (Analysis 1). Most important for our hypotheses, there was strong evidence against the 3-way interaction of set size, PI, and distraction (Experiment 1: $BF_{01} = 82.6$; Experiment 3: $BF_{01} = 54.2$). In both experiments there was strong support for the interaction of set size and PI. In addition, distraction had a detrimental main effect on performance. The interaction of distraction with set size reflects a stronger detrimental effect of distraction at the smaller set size, as would be expected if performance with the smaller set size relied more exclusively on WM, whereas performance with the larger set size relied in part on eLTM, and WM – but not eLTM – is vulnerable to distraction by a secondary task (Glanzer & Cunitz, 1966).

Across both set sizes together there was no evidence for a main effect of PI. To assess whether there was PI at the larger set size, we ran follow-up analyses focusing on set size 8 only (Analysis 2). For that set size, there was strong evidence for the main effect of PI, and against its interaction with distraction condition (see Tables 2 and 3 for the corresponding BF_{10} values) .

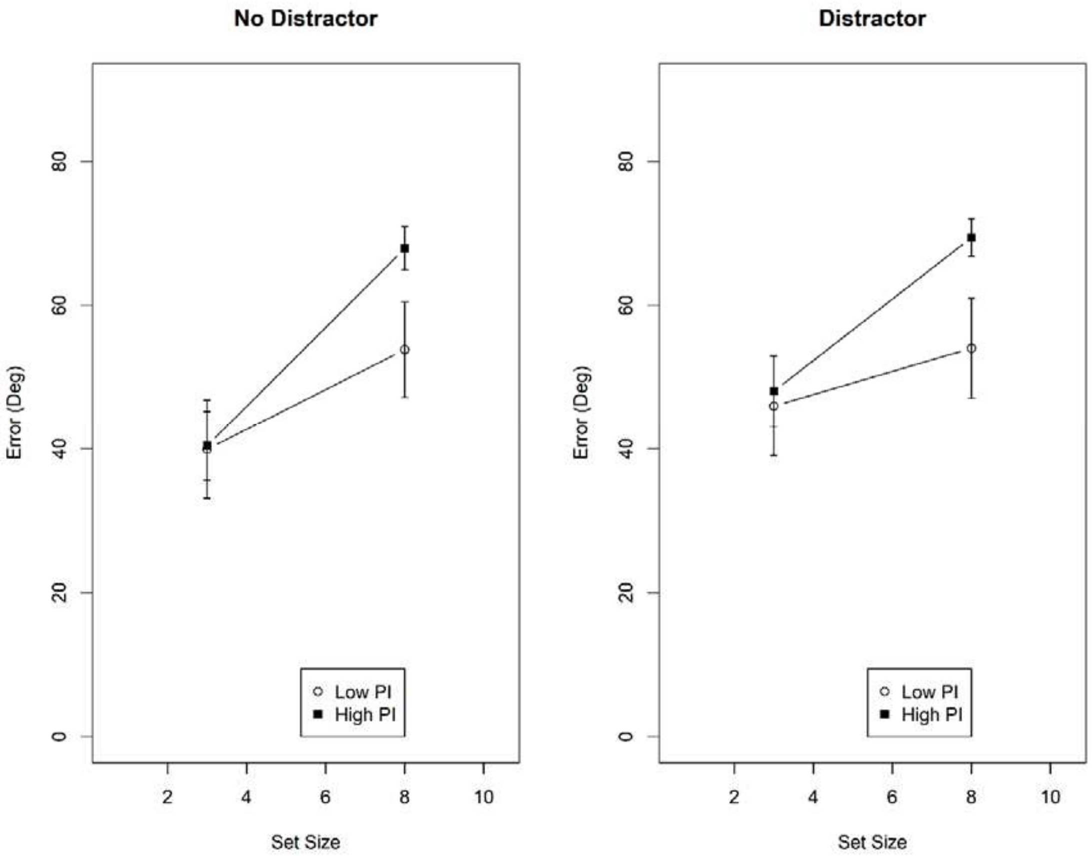

Figure 5: Experiment 2: Mean error of reproduction as a function of set size and PI, for conditions without and with distraction by one subitizing trial in the retention interval.

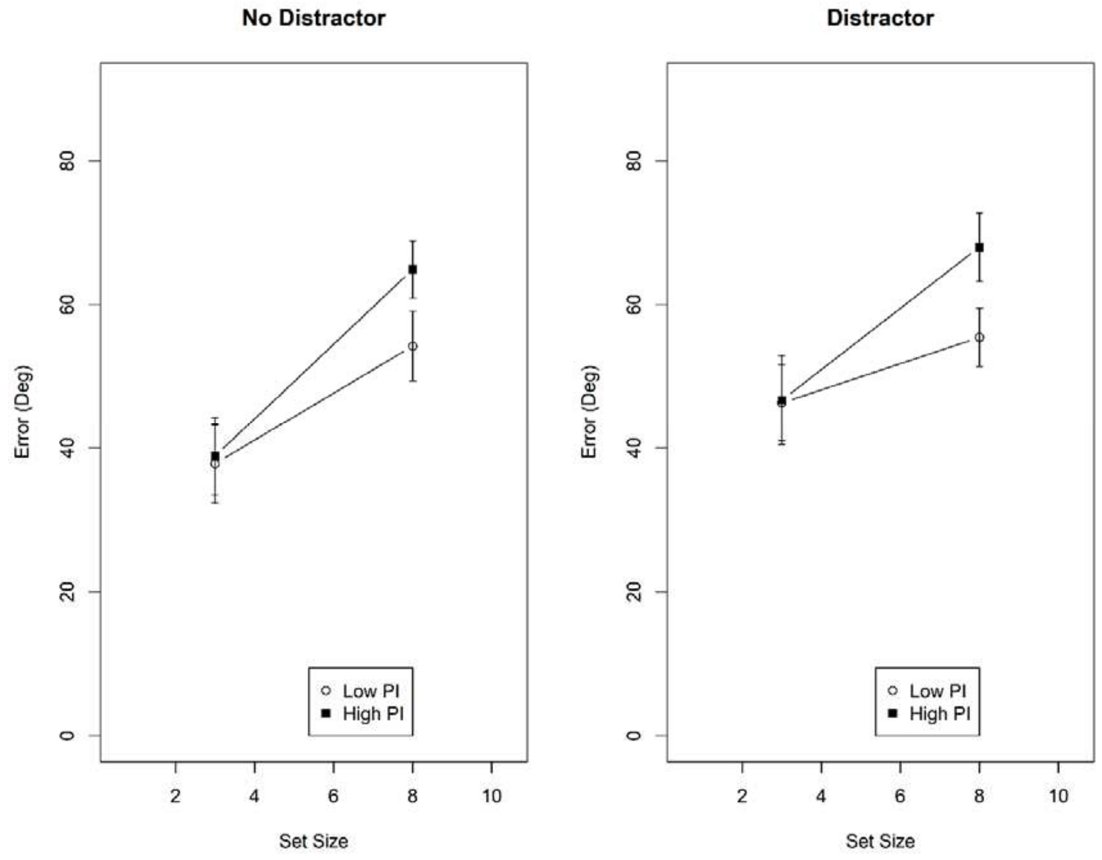

Figure 6: Experiment 3: Mean error of reproduction as a function of set size and PI, for conditions without and with distraction by one subitizing trial in the retention interval.

Table 2: Bayes Factors (BF_{10}) in favor of the fixed effects, Experiment 2 .

Analysis 1	**Main Effect**	**Interaction with PI**	**Interaction with Distraction**
Set size	5.05×10^{13}	1666	8.0
PI	0.6		0.016
Distraction	670		
Analysis 2 (Set size 8 only)			
PI	53.2		0.017
Distraction	0.021		

Table 3: Bayes Factors (BF_{10}) in favor of the fixed effects, Experiment 3.

Analysis 1	**Main Effect**	**Interaction with PI**	**Interaction with Distraction**	**Interaction with Delay / Number of Distractors**
Set size	5.2×10^{21}	2163	68.0	
PI	0.28		0.01	
Distraction	8.7×10^{8}			
Analysis 2 (Set size 8 only)				
PI	25.5		0.024	
Distraction	0.49			
Analysis 3				
Set size	5.2×10^{21}	2205	69.2	0.01
PI	0.27		0.015	1.15
Distraction	8.9×10^{8}			0.03
Length of Delay / Number of Distractors	0.017			

With Experiment 3 we tested again whether PI emerged at the smaller set size when the retention interval was prolonged (in the no-distraction condition), or when more distractor trials had to be carried out, in addition to a longer retention interval (in the distraction condition). Figure 7 shows that this was not the case. Even after a 12 s delay and 4 subitizing trials, PI was observed only for set size 8, not for set size 3. The statistical evaluation of this pattern (Analysis 3) relied on a model with 4 predictors: Set size, PI, distraction condition, and delay duration (including the number of distractors, if any); here we focus only on the new effects involving the latter predictor. There was compelling evidence against the four-way interaction, BF_{01} = 70.0, as well as all three-way interaction involving delay duration (set size x PI x delay: BF_{01} = 97.8; set size x distraction x delay: BF_{01} = 64.5; PI x distraction x delay, BF_{01} = 105.7). A model omitting the four-way and all three-way interactions received unambiguous support compared to the model with all fixed effects, $BF_{10} = 2.6 \times 10^9$. Therefore, we conclude that delay duration does not modulate the interaction of set size and PI. The subsequent tests of lower-order effects took the model with all main effects and two-way interactions for comparison; Table 3 summarizes the BF_{10}. There was strong evidence against the interaction of delay duration with set size, (BF_{10} = 0.105; BF_{01} = 95.2), and of delay duration with distraction condition, (BF_{10} = 0.032; BF_{01} = 31.0); evidence on the interaction with PI was ambiguous, BF_{10} = 1.15. Finally, there was strong evidence against a main effect of delay duration, (BF_{10} = 0.017; BF_{01} = 59.0).

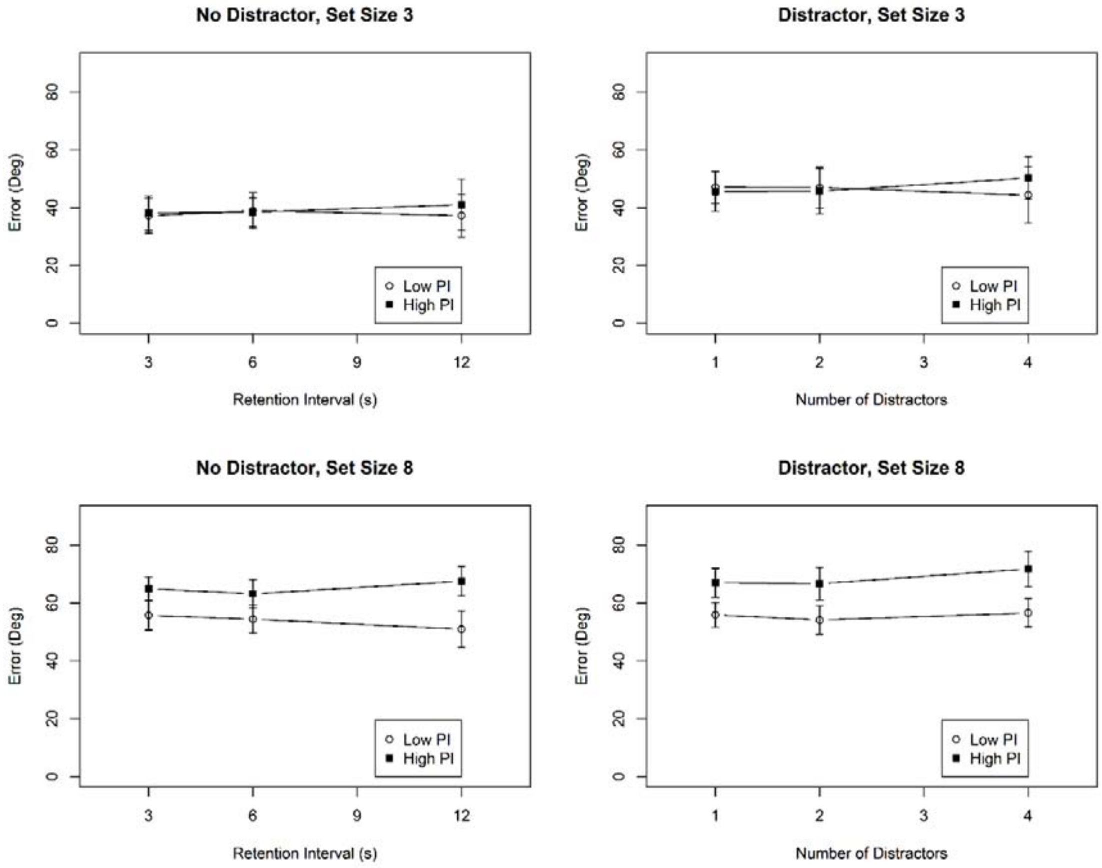

Figure 6: Experiment 3: Mean error of reproduction as a function of the retention interval (no-distraction condition) and as a number of distractor trials (distraction condition).

A further analysis on set size 3 only revealed strong evidence against a main effect of PI, (BF_{10} = 0.059; BF_{01} = 17.0, against the interactions of PI with distraction condition, (BF_{10} = 0.019; BF_{01} = 53.4) and the interaction of PI with delay (BF_{10} = 0.047; BF_{01} = 21.1). Therefore, we can be confident that PI does not arise with longer retention intervals when the set size is within WM capacity.

Discussion

Experiments 2 and 3 provide evidence for an activity-silent form of WM. With small set sizes that did not exceed WM capacity, subjects were able to reproduce the memoranda without any evidence of PI, even after an interruption of storage-related neural activity. The PI-resistance with smaller set sizes contrasts with robust evidence of PI with larger set sizes that are expected to garner contributions from eLTM, a finding that demonstrates our task's sensitivity to this key signature of eLTM function. The critical interaction of PI with set size does not depend on the duration of the retention interval, confirming the conclusion from Experiment 1 that PI at the larger set sizes is not caused by reduced temporal distinctiveness. The number of distractors did not modulate that pattern either, and it did not even have a main effect. Apparently a single subitizing trial is enough to cause the – modest – impairment of WM performance, and additional trials do not add to that damage. One explanation for that pattern is that a single subitizing trial wipes out the neural activity that is reflected in the CDA, but does not affect the representation in the activity-silent form of WM. Additional subitizing trials cause no further damage because there is no neural activity left to disrupt.

Experiment 4

With the final experiment we focus on a second neural signature of maintenance in WM, the decodability of memory contents from the spatial pattern of EEG alpha power. When people hold one or two locations on a circle in WM, these locations can be decoded with an inverted-encoding model from the alpha-power distribution over electrodes (Foster et al., 2016; Sutterer, Foster, Adam, Vogel, & Awh, 2019). This decodability is temporarily disrupted by a single visual-search trial during the retention interval, with only negligible damage to memory performance (van Moorselaar et al., 2017). Here we investigate whether the same visual-search task forces participants to outsource memory for locations to eLTM, so that it suffers PI even at small set sizes.

Method

Participants. Sixty students of the University of Zurich took part in exchange for partial course credit or 15 Swiss Francs. They were assigned at random to two groups.

Design, Materials, and Procedure. The design was identical to Experiment 2. The memory task was adapted so that we could test memory for location: The same objects as in the preceding experiments were displayed, but this time in white, and on a randomly chosen location on a virtual circle centered on the screen center, with a radius of 0.2 times the screen width (about 6 degrees of visual angle at a distance of 1 m). At test, instead of a color wheel, participants saw a dark grey circular band with the same radius as the object locations, and they were instructed to select the location of the test object in that band. In all other respects the memory task was the same as in Experiment 2.

Instead of subitizing, the distractor task was modeled after the visual search task of van Moorselaar et al. (2017). After the offset of the last memory stimulus, a 200 ms blank screen, and a fixation cross shown for 300 ms, participants saw a circular arrangement of eight colored frames; the 2 frames in opposite locations had the same color. After 800 ms, each frame was filled with a character. Participants were instructed to determine whether one of them was a digit, or all of them were letters. In 87.5 % of the trials in which a digit appeared, it appeared in one of the two red frames, and participants were informed of that. The characters were masked by the digit “8” after a short interval, determined individually through calibration to approximate 66% correct responses (Kaernbach, 1991) in the first phase of the experiment, in which participants did 64 search trials. In the second phase, when the search task was combined with the memory task, the presentation duration of the search stimuli was held constant at the calibrated time for each participant. Participants completed 6 test blocks of the memory task, each with 16 trials (8 per set size).

Results

The results, presented in Figure 7, replicate the key finding of the preceding two experiments: PI interacted with set size, and that interaction did not depend on whether or not there was a distraction task in the retention interval: There was strong evidence against the three-way interaction, BF_{01} = 61.3. The BFs of the remaining effects are summarized in Table 4.

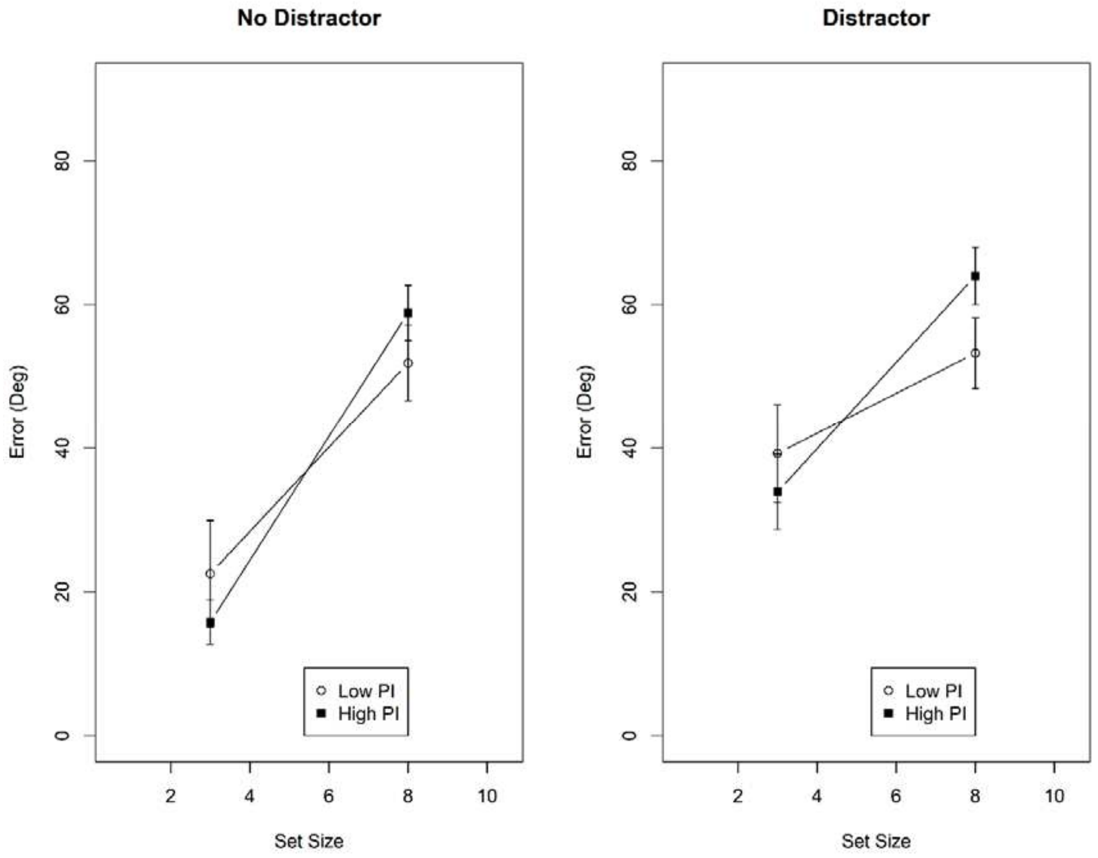

Figure 7: Experiment 4: Mean error of reproduction as a function of set size and PI, for conditions without and with distraction by one visual-search trial in the retention interval.

The pattern of effects in Figure 7 appears to differ from that in the preceding experiments in that there was at best a small detrimental effect of PI at set size 8. Indeed, evidence for the main effect of PI at set size 8 was only ambiguous, BF_{10} = 1.6. The interaction of PI with set size was in part due to the apparent slight beneficial effect of high PI at set size 3 (although the main effect of PI at set size 3 was not statistically supported, BF_{10} = 0.18). Compared to Experiments 2 and 3, it looks as if there was an overall beneficial effect of the condition with high PI independent of set size, together with the detrimental effect only at the larger set size. We can only speculate about what might have caused the beneficial effect: In the condition with high PI participants acquire more familiarity with the often-repeated objects, and that could have facilitated binding these objects to locations more than facilitating their bindings to colors. Evidence that familiarity of visual objects facilitates memory for bindings involving these objects, at least in eLTM, comes from a study by Reder, Liu, Keinath, and Popov (2016).

Additional evidence for a contribution of eLTM to location recall comes from the analysis of intrusions from previous trials (Supplementary, https://osf.io/tp49k/).

Table 4: Bayes Factors (BF_{10}) in favor of the fixed effects, Experiment 4.

Experiment 4	**Main Effect**	**Interaction with PI**	**Interaction with Distraction**
Set size	2.0×10^{30}	3.0×10^{5}	1.4×10^{27}
PI	0.06		0.41
Distraction	8.1×10^{57}		

Discussion

Experiment 4 extends our result to a procedure that past work has used to measure spatially-selective alpha activity during a spatial WM task. This alpha signature of WM storage was temporarily eliminated by the imposition of a visual search task during the delay period (van Morselaar et al., 2018). Here, the same distractor task did not lead to PI when the memory set size was within WM capacity. Only with a set size clearly exceeding WM capacity was there a hint of PI, regardless of distraction. This is the result expected on the assumption that a small set of spatial locations can be maintained in a form of WM that is not accompanied by spatially-selective patterns of alpha activity.

General Discussion

When neural signatures of maintenance of information in WM are disrupted by withdrawing attention from it, people's ability to remember the information is often preserved. This ability could reflect a form of "activity silent" WM, or eLTM. Across three experiments we tested competing predictions from these two explanations, using PI-resistance as the functional "fingerprints" of WM storage. The results consistently favored the assumption of "activity silent" WM: After distraction by a secondary task, memory sets within the capacity of WM are still maintained without PI from previous trials. This contrasts with memory sets clearly exceeding WM capacity, for which we observed PI regardless of whether or not people were distracted.

Taken together, these findings support the following conclusions: In experimental paradigms typically used for investigating WM, representations are formed both in WM and in eLTM, though apparently more slowly in the latter. Whereas eLTM is prone to PI from outdated information, such as stimuli from previous trials, that information is cleared from WM, so that information in WM does not suffer interference from it. When possible, people meet the memory demand at test by drawing only on information in WM, but when that fails – for instance, when the memory set far exceeds WM capacity – they draw on eLTM as a backup. When people are forced to attend to and process distracting information during the retention interval, that impairs, but does not abolish, their ability to hold the memoranda in WM. This is the case even if the distraction eliminates neural signals that correlate with maintenance in WM, such as the CDA, and patterns of alpha activity that track the spatial position of stored items.

The notion of WM maintenance in the absence of known neural correlates has been referred to as “activity silent” WM (Stokes, 2015). This label should be interpreted cautiously: Our conclusion that there is a form of “activity silent” WM means that there is a form of temporary memory that meets defining characteristics of WM – a limited capacity, and protection against PI from outdated information – and functions despite the interruption of specific established neural correlates of WM maintenance. As it is not possible to measure the absence of all possible kinds of neural activity, we do not wish to claim that this form of WM is accompanied by no neural activity at all (cf. Christophel, Iamshchinina, Yan, Allefeld, & Haynes, 2018)

If neural processes reflected in the CDA, and in the distribution of alpha power, are not necessary for maintenance in WM, then what is their function? One possibility is that there are two kinds of WM, one reliant on these neural processes and the other not. In the absence of disruption people can draw on both to guide their behavior; with distraction they lose the first, leading to impaired but not catastrophic performance. Another possibility is that these neural processes underlie attention to the contents of WM. This interpretation is supported by the fact that both the CDA and the decodability of spatial information from alpha power have been observed in tasks that require attention to, but no memory of, the stimuli. For instance, the CDA is elicited by subitizing (Mazza & Caramazza, 2015) and multi-object tracking (Drew, Horowitz, Wolfe, & Vogel, 2011; Drew & Vogel, 2008). Decoding of locations from alpha power has been demonstrated for locations that people attend to without having to remember them (Foster, Sutterer, Serences, Vogel, & Awh, 2017). This interpretation has the advantage of parsimony because it does not require introducing a new construct – a second form of WM. With this interpretation, the conclusion from the present work can be described as as additional confirmation of an assumption that other findings have already supported (Oberauer, 2019): Information can be maintained in WM without persistent attention to it.

One could object to this interpretation that WM, by definition, involves directing attention towards items that are held “in mind”. Indeed, past theorists have often posited an equivalence between WM and the “focus of attention” (e.g., Cowan, 2001; Foster et al., in press; Hakim et al., 2021). Although we acknowledge that the memories were not held “in mind” during the interruption by a secondary task, we have provided clear positive evidence that these activity-silent representations share two key functional signatures of WM storage: The representations are protected from PI, and they are subject to a similarly limited capacity. Thus, by revealing novel functional characteristics of these activity-silent representations, our work provides an important complement to past demonstrations that memory representations can still be accessed following interruptions of storage-related patterns of neural activity. Future research might analyze how they differ functionally (Masse, Yang, Song, Wang, & Freedman, 2019): Which cognitive process can only operate with a representation when it is attended to, and accompanied by a characteristic signature of neural activity?

References

Beukers, A. O., Buschman, T. J., Cohen, J. D., & Norman, K. A. (2021). Is Activity Silent Working Memory Simply Episodic Memory? *Trends in Cognitive Sciences, 25*(4), 284-293. doi:https://doi.org/10.1016/j.tics.2021.01.003

Botvinick, M., & Watanabe, T. (2007). From numerosity to ordinal rank: A gain-field model of serial order representation in cortical working memory. *Journal of Neuroscience, 27*, 8636-8642.

Brady, T. F., Konkle, T. F., Gill, J., Oliva, A., & Alvarez, G. A. (2013). Visual long-term memory has the same limit on fidelity as visual working memory. *Psychological Science, 24*, 981-990. doi:10.1177/0956797612465439

Brown, G. D. A., Neath, I., & Chater, N. (2007). A temporal ratio model of memory. *Psychological Review, 114*, 539-576.

Burgess, N., & Hitch, G. J. (1999). Memory for serial order: A network model of the phonological loop and its timing. *Psychological Review, 106*, 551-581.

Bürkner, P.-C. (2017). brms: An R package for Bayesian multilevel models using Stan. *Journal of Statistical Software, 80*, 1-28. doi:10.18637/jss.v080.i01

Christophel, T. B., Hebart, M. N., & Haynes, J.-D. (2012). Decoding the contents of visual short-term memory from human visual and parietal cortex. *Journal of Neuroscience, 32*, 12983-12989. doi:10.1523/JNEUROSCI.0184-12.2012

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

Cowan, N. (2001). The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behavioral and Brain Sciences, 24*, 87-185.

Drew, T., Horowitz, T. S., Wolfe, J. M., & Vogel, E. K. (2011). Delineating the Neural Signatures of Tracking Spatial Position and Working Memory during Attentive Tracking. *The Journal of Neuroscience, 31*(2), 659. doi:10.1523/JNEUROSCI.1339-10.2011

Drew, T., & Vogel, E. K. (2008). Neural measures of individual differences in selecting and tracking multiple moving objects. *The Journal of Neuroscience, 28*(16), 4183.

Foster, J. J., Sutterer, D. W., Serences, J. T., Vogel, E. K., & Awh, E. (2016). The topography of alpha-band activity tracks the content of spatial working memory. *Journal of Neurophysiology, 115*, 168-177. doi:10.1152/jn.00860.2015.

Foster, J. J., Sutterer, D. W., Serences, J. T., Vogel, E. K., & Awh, E. (2017). Alpha-Band Oscillations Enable Spatially and Temporally Resolved Tracking of Covert Spatial Attention. *Psychological Science, 28*(7), 929-941. doi:10.1177/0956797617699167

Foster, J. J., Vogel, E. K., & Awh, E. (in press). Working memory as persistent neural activity *Oxford Handbook of Human Memory*.

Glanzer, M., & Cunitz, A. R. (1966). Two storage mechanisms in free recall. *Journal of Verbal Learning & Verbal Behavior, 5*, 351-360.

Gronau, Q. F., Singmann, H., & Wagenmakers, E. J. (2018). *bridgesampling: An R package for estimating normalization constants*. arXiv, https://arxiv.org/pdf/1710.08162.pdf.

Hakim, N., Feldmann-Wüstefeld, T., Awh, E., & Vogel, E. K. (2019). Perturbing Neural Representations of Working Memory with Task-irrelevant Interruption. *Journal of Cognitive Neuroscience, 32*(3), 558-569. doi:10.1162/jocn_a_01481

Hakim, N., Feldmann-Wüstefeld, T., Awh, E., & Vogel, E. K. (2021). Controlling the Flow of Distracting Information in Working Memory. *Cerebral Cortex, 31*(7), 3323-3337. doi:10.1093/cercor/bhab013

Harrison, S., & Tong, F. (2009). Decoding reveals the contents of visual working memory in early visual areas. *Nature, 458*, 632-635. doi:10.1038/nature07832

Kaernbach, C. (1991). Simple adaptive testing with the weighted up-down method. *Perception & Psychophysics, 49*, 227-229.

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*(1), 127-142. doi:10.1162/jocn_a_00305

Lewandowsky, S., & Farrell, S. (2008). Short-term memory: new data and a model. In B. H. Ross (Ed.), *The Psychology of Learning and Motivation* (Vol. 49, pp. 1-48). London, UK: Elsevier.

Lewis-Peacock, J. A., Drysdale, A. T., Oberauer, K., & Postle, B. R. (2011). Neural evidence for a distinction between short-term memory and the focus of attention. *Journal of Cognitive Neuroscience, 24*, 61-79.

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

Masse, N. Y., Yang, G. R., Song, H. F., Wang, X.-J., & Freedman, D. J. (2019). Circuit mechanisms for the maintenance and manipulation of information in working memory. *Nature Neuroscience, 22*, 1159–1167. doi:10.1038/s41593-019-0414-3

Mazza, V., & Caramazza, A. (2015). Multiple object individuation and subitizing in enumeration: a view from electrophysiology. *frontiers in human neuroscience, 9*. Retrieved from doi:10.3389/fnhum.2015.00162

Morey, R. D., & Rouder, J. N. (2015). BayesFactor (Version 0.9.12.2). Retrieved from http://cran.at.r-project.org/web/packages/BayesFactor/index.html

Oberauer, K. (2019). Working memory and attention - a conceptual analysis and review. *Journal of Cognition, 2*, 1-23. doi:10.5334/joc.58

Oberauer, K. (2022). The Importance of Random Slopes in Mixed Models for Bayesian Hypothesis Testing. *Psychological Science, 33*(4), 648-665. doi:10.1177/09567976211046884

Oberauer, K., Awh, E., & Sutterer, D. W. (2017). The role of long-term memory in a test of visual working memory: proactive facilitation but no proactive interference. *Journal of Experimental Psychology: Learning, Memory and Cognition, 43*, 1-22.

Oberauer, K., & Greve, W. (2021). Intentional remembering and intentional forgetting in working and long-term memory. *Journal of Experimental Psychology: General*. doi:10.1037/xge0001106

Oberauer, K., Lewandowsky, S., Farrell, S., Jarrold, C., & Greaves, M. (2012). Modeling working memory: An interference model of complex span. *Psychonomic Bulletin & Review, 19*, 779-819. doi:10.3758/s13423-012-0272-4

Oberauer, K., & Lin, H.-Y. (2017). An interference model of visual working memory. *Psychological Review, 124*, 21-59.

Reder, L. M., Liu, X. L., Keinath, A., & Popov, V. (2016). Building knowledge requires bricks, not sand: The critical role of familiar constituents in learning. *Psychonomic Bulletin & Review, 23*(1), 271-277. doi:10.3758/s13423-015-0889-1

Rose, N. S., LaRocque, J. J., Riggall, A. C., Gosseries, O., Starrett, M. J., Meyering, E. E., & Postle, B. R. (2016). Reactivation of latent working memories with transcranial magnetic stimulation. *Science, 354*, 1136-1139. doi:10.1126/science.aah7011

Rouder, J. N., Morey, R. D., Speckman, P. L., & Province, J. M. (2012). Default Bayes factors for ANOVA designs. *Journal of Mathematical Psychology, 56*, 356-374.

Serences, J. T., Ester, E. F., Vogel, E. K., & Awh, E. (2009). Stimulus-specific delay activity in human primary visual cortex. *Psychological Science, 20*, 207-214.

Singmann, H., Kellen, D., Cox, G. E., Chandramouli, S. H., Davis-Stober, C. P., Dunn, J. C., . . . Shiffrin, R. M. (2021). *Statistics in the service of science: Don't let the tail wag the dog*. (DOI: 10.31234/osf.io/kxhfu). PsyArXiv.

Sprague, T. C., Ester, E. F., & Serences, J. T. (2016). Restoring latent visual working memory representations in human cortex. *Neuron, 91*, 694-707. doi:10.1016/j.neuron.2016.07.006

Stokes, M. G. (2015). 'Activity-silent' working memory in prefrontal cortex: a dynamic coding framework. *Trends in Cognitive Sciences, 19*, 394-405. doi:10.1016/j.tics.2015.05.004

Sutterer, D. W., Foster, J. J., Adam, K. C. S., Vogel, E. K., & Awh, E. (2019). Item-specific delay activity demonstrates concurrent storage of multiple active neural representations in working memory. *PLoS Biology, 17*(4), e3000239. doi:10.1371/journal.pbio.3000239

van Moorselaar, D., Foster, J. J., Sutterer, D. W., Theeuwes, J., Olivers, C. N. L., & Awh, E. (2017). Spatially selective alpha oscillations reveal moment-by-moment trade-offs between working memory and attention. *Journal of Cognitive Neuroscience, 30*, 256-266. doi:10.1162/jocn_a_01198

Vogel, E. K., McCollough, A. W., & Machizawa, M. G. (2005). Neural measures reveal individual differences in controlling access to working memory. *Nature, 438*(24), 500-503.

Wolff, M. J., Ding, J., Myers, N. E., & Stokes, M. G. (2015). Revealing hidden states in visual working memory using electroencephalography. *frontiers in Systems Neuroscience, 9*. Retrieved from doi:10.3389/fnsys.2015.00123