

Journal of
Cognitive Neuroscience**Perturbing neural representations of working memory with
task-irrelevant interruption**

Journal:	<i>Journal of Cognitive Neuroscience</i>
Manuscript ID	JOCN-2019-0310
Manuscript Type:	Original
Date Submitted by the Author:	17-Sep-2019
Complete List of Authors:	Hakim, Nicole; University of Chicago, Psychology; University of Chicago Feldmann-Wüstefeld, Tobias; University of Southampton, Psychology Awh, Edward; University of Chicago, Vogel, Edward; University of Chicago, Department of Psychology
Keywords:	Memory: Working memory, Attention: Visual, interruption, Electrophysiology

Perturbing neural representations of working memory with task-irrelevant interruption

Nicole Hakim^{1,2,*}, Tobias Feldmann-Wüstefeld³, Edward Awh^{1,2,4} & Edward K. Vogel^{1,2,4}

1. Department of Psychology, University of Chicago, Chicago, Il. 606237
2. Institute for Mind and Biology, University of Chicago, Chicago, Il. 606237
3. Department of Psychology, University of Southampton, Southampton, U.K.
SO17 1BJ
4. Grossman Institute for Neuroscience, Quantitative Biology, and Human Behavior, University of Chicago, Chicago, Il. 606237

* Corresponding author

Abbreviated title: INTERRUPTION OF ONLINE REPRESENTATIONS

Key words: visual working memory, attention, interruption, electrophysiology

Contributions: All authors conceived and designed experiments and drafted and revised the manuscript. N.H. and T.F.W. performed analyses. N.H. collected the data.

Acknowledgements: Research was supported by NIMH grant 5ROI MH087214-08 and Office of Naval Research grant N00014-12-1-0972.

Data Availability: Datasets for all experiments will become available online on Open Science Framework upon acceptance of the manuscript or reviewer request.

Conflicts of Interest: none

Correspondence to:

Nicole Hakim

University of Chicago

940 East 57th Street, Chicago, IL 60637, USA

nhakim@uchicago.edu, 773-795-4784

INTERRUPTION OF ONLINE REPRESENTATIONS

2

Abstract

Working memory maintains information so that it can be used in complex cognitive tasks. A key challenge for this system is to maintain relevant information in the face of task-irrelevant perturbations. Across two experiments, we investigated the impact of task-irrelevant interruptions on neural representations of working memory. We recorded electroencephalogram (EEG) activity in humans while they performed a working memory task. On a subset of trials, we interrupted participants with salient, but task-irrelevant objects. To track the impact of these task-irrelevant interruptions on neural representations of working memory, we measured two well-characterized, temporally sensitive EEG markers that reflect active, prioritized working memory representations: the contralateral delay activity (CDA) and lateralized alpha power (8-12hz). Following interruption, we found that CDA amplitude momentarily sustained, but was gone by the end of the trial. Lateralized alpha power was immediately influenced by the interrupters, but recovered by the end of the trial. This suggests that dissociable neural processes contribute to the maintenance of working memory information, and that brief irrelevant onsets disrupt two distinct online aspects of working memory. Additionally, we found that task expectancy modulated the timing and magnitude of how these two neural signals responded to task-irrelevant interruptions, suggesting that the brain's response to task-irrelevant interruption is shaped by task context.

INTERRUPTION OF ONLINE REPRESENTATIONS

3

Introduction

Working memory is a large-scale neural system that maintains readily accessible task-relevant information via active neural firing. A key challenge for this system is to protect these active representations from task-irrelevant interruptions. Extensive prior work has characterized how the presence of irrelevant information during the encoding of targets (distractors) impacts working memory representations (Postle, D'Esposito and Corkin, 2005; Clapp, Rubens and Gazzaley, 2010; Gaspar and McDonald, 2014; Feldmann-Wüstefeld and Vogel, 2018). This work has revealed that the presence of distractors during this initial encoding period (0-500ms) greatly reduces working memory performance, in part because these items compete with targets for limited representational space in working memory (De Fockert *et al.*, 2001; Vogel, McCollough and Machizawa, 2005; Olivers, 2008). After this initial encoding period, presence of irrelevant information (interrupters) has a reduced, but still measurable impact on performance (Vogel, Woodman and Luck, 2006). These interrupters have less of an impact because working memory representations have reached a more stable state, which is consistent with the time-course of neural measures of working memory representations (Vogel, Woodman and Luck, 2006; Ikkai, McCollough and Vogel, 2010). This reduced impact is also likely due to the formation of concurrent visual long-term memory representations that represent the targets in a passive, yet still accessible format (e.g. Fukuda & Vogel, in press; Woodman and Chun, 2006; Chun and Turk-Browne, 2007). Together, these concurrent active and passive representations of targets reduce the behavioral impact of interruption during working memory maintenance. Yet, despite robust behavioral performance, current models of working memory still predict that onsets of task-irrelevant interruption should produce a momentary perturbation of the maintained target representations during which attention is withdrawn from the targets and at least temporarily applied to the positions of the interrupters (Bisley and Goldberg, 2003; Bisley *et al.*, 2004). However, the consequences of such a brief withdrawal of attention on the neural signatures of working memory are not well understood. Here, we seek to measure the impact that task-irrelevant interruption has on the ongoing active neural representations of targets held in working memory.

To track the impact of task-irrelevant interruption on neural representations of working memory, we measured two well-characterized, temporally sensitive EEG markers

INTERRUPTION OF ONLINE REPRESENTATIONS

4

that reflect active, prioritized working memory representations: the contralateral delay activity (CDA) and lateralized alpha power (8-12hz). The CDA is a sustained negative-going wave in human EEG that tracks current working memory load. It is sensitive to trial-by-trial fluctuations in working memory performance and distinguishes stable individual differences in working memory (Vogel and Machizawa, 2004; Luria *et al.*, 2016). This component is thought to reflect an index of the current items that are actively represented in working memory (Feldmann-Wüstefeld, Vogel and Awh, 2018; Hakim *et al.*, 2018). Lateralized alpha power is similarly sensitive to task-relevant information. This signal is measured as a decrease in alpha power over posterior electrodes that are contralateral to the position of the attended items. However, despite its similarity to the CDA, it has been shown to be a distinct component of actively maintained information (Fukuda, Mance and Vogel, 2015) that appears to primarily track the current position of spatial attention (Worden *et al.*, 2000; Thut *et al.*, 2006; Foster *et al.*, 2016, 2017). Topographic distributions of alpha power across the entire scalp have been shown to contain precise spatial information about remembered/attended stimuli (Foster *et al.*, 2017; van Moorselaar *et al.*, 2017), whereas lateralized alpha power have been used as an effective tool for establishing which visual hemifield is currently attended. Together, the CDA and lateralized alpha power respectively provide an item-based and space-based index of task-relevant information that is actively represented in working memory. Furthermore, because both signals are lateralized, we were able to isolate processing of the memory array by presenting the memory items laterally and the interrupters along the vertical midline of the display. As items on the vertical midline do not affect lateralized signals, the ongoing lateral measures only reflect processing of the memory representations. This allowed us to measure how these working memory representations respond to task-irrelevant interruption.

Materials & Methods

Overview of Experiments

In Experiment 1, we sought to determine how task-irrelevant interrupters impact ongoing working memory representations. We did this by presenting midline interrupters during the retention interval of a working memory task. In Experiment 2, we sought to

INTERRUPTION OF ONLINE REPRESENTATIONS

5

determine whether the neural responses to task-irrelevant interrupters could be modulated by task expectancy. During all of these tasks, we recorded EEG activity from human participants.

Experiment 1

Participants. Twenty-two volunteers, naïve to the objective of the experiment participated for payment (~15 USD per hour). All data were collected in a single session. The data of 2 participants were excluded from the analysis because of too many artifacts, poor behavioral performance (see below for criteria) or technical problems. The remaining 20 participants (12 male) were between the ages of 19-30 ($M = 22.7$, $SD = 3.4$). Participants in all experiments reported normal or corrected-to-normal visual acuity as well as normal color vision. All experiments were conducted with the written understanding and consent of each participant. The University of Chicago Institutional Review Board approved experimental procedures.

Stimuli. All stimuli were presented on a gray background (~33.3 cd/m^2). Cue displays showed a small central fixation dot ($0.2^\circ \times 0.2^\circ$). A horizontal diamond comprised of a green (RGB = 74, 183, 72; 52.8 cd/m^2) and a pink (RGB = 183, 73, 177; 31.7 cd/m^2) triangle appeared on the vertical midline 0.65° above the fixation dot. In 50% of the trials, the pink triangle pointed to the left side and the green triangle pointed to the right side, in the remaining 50% of the trials this was inverse. Half the participants were instructed to attend the hemifield that the pink triangle pointed to, and the other half was instructed to attend the hemifield the green triangle pointed to. Memory displays showed a series of colored squares (1.1° by 1.1° , mean luminance 43.1 cd/m^2). Colors for the squares were selected randomly from a set of 11 possible colors (Red = 255 0 0; Green = 0 255 0; Blue = 0 0 255; Yellow = 255 255 0; Magenta = 255 0 255; Cyan = 0 255 255; Purple = 102 0 102; Brown = 102 51 0; Orange = 255 128 0; White = 255 255 255; Black = 0 0 0). Squares could appear within an area of the display subtending 3.5° to the left or right of fixation and 3.1° above and below fixation. There was the same number of squares in each hemisphere. Within each hemisphere, squares were as equally distributed between the upper and lower hemifields as possible. The interruption display showed four colored squares of the same size as the memory display, drawn from the remaining colors. These interrupting items were shown on the vertical midline with a randomly jittered horizontal

INTERRUPTION OF ONLINE REPRESENTATIONS

6

offset of maximally 0.55° (half of an object). Retention interval displays were blank with a small central fixation dot ($0.2^\circ \times 0.2^\circ$). Probe displays showed one colored square in each hemisphere in the same location as one of the squares in the original array. In 50% of the trials, the color was identical (no change trials) to the memory display. In the remaining 50% of the trials, it was one of the colors not used in the memory or interruption display (change trials). The same stimuli were used in all experiments.

Apparatus. Participants were seated with a chin-rest in a comfortable chair in a dimly lit, electrically shielded and sound attenuated chamber. Participants responded with button presses on a standard keyboard that was placed in front of them. Stimuli were presented on an LCD computer screen (BenQ XL2430T; 120 Hz refresh rate; 61 cm screen size in diameter; 1920×1080 pixels) placed at 74 cm distance from participants. An IBM-compatible computer (Dell Optiplex 9020) controlled stimulus presentation and response collection.

Procedure. Each trial began with a cue display (500 ms) indicating the relevant side of the screen (left or right). A memory display consisting of 6 colored squares in each hemisphere followed the cue display for 150 ms. Participants were instructed to memorize as many colored squares in the memory display from the cued side and to ignore the other side entirely, as that side would never be probed. Participants had to remember the items for a retention interval of 1,650 ms during which a central fixation dot was shown. In 25% of the trials, an interruption display appeared 500 ms after memory display offset for 150 ms. The total length of the retention interval was 1,650 ms, regardless of whether an interruption appeared. Participants were instructed to always ignore interruption displays. After the retention interval, a probe display appeared until response. Participants had to indicate whether the color at the probed location changed color (“?” key) or did not change color (“z” key). After participants responded, the trial concluded and the next trial started after a blank inter-trial interval of 750 ms. Participants completed a total of 1200 trials (15 blocks of 80 trials), i.e. 300 trials with interruption and 900 trials without interruption. Information about average performance and a minimum break of 30 seconds was provided after each block. See Figure 1 for a visual depiction of the task.

We presented the interrupters in locations that did not overlap with the locations of the memory items to avoid visual masking. Importantly, the relative position of interrupters

INTERRUPTION OF ONLINE REPRESENTATIONS

7

and targets matter in lateralized change detection tasks. When interrupters are presented laterally with targets on the vertical midline, the neural signature of sustained interrupters suppression can be isolated (CDAp). Conversely, when interrupters are presented on the vertical midline and targets are presented laterally, the neural signature of target processing can be isolated (Feldmann-Wüstefeld & Vogel, 2018). Accordingly, since we were interested in how neural representations of targets are affected by interruption, we placed the interrupters along the vertical midline. Thus, reductions in CDA amplitude can be interpreted as dropping memory items, and reductions in lateralized alpha power can be interpreted as a shift of attention away from the laterally presented memory arrays.

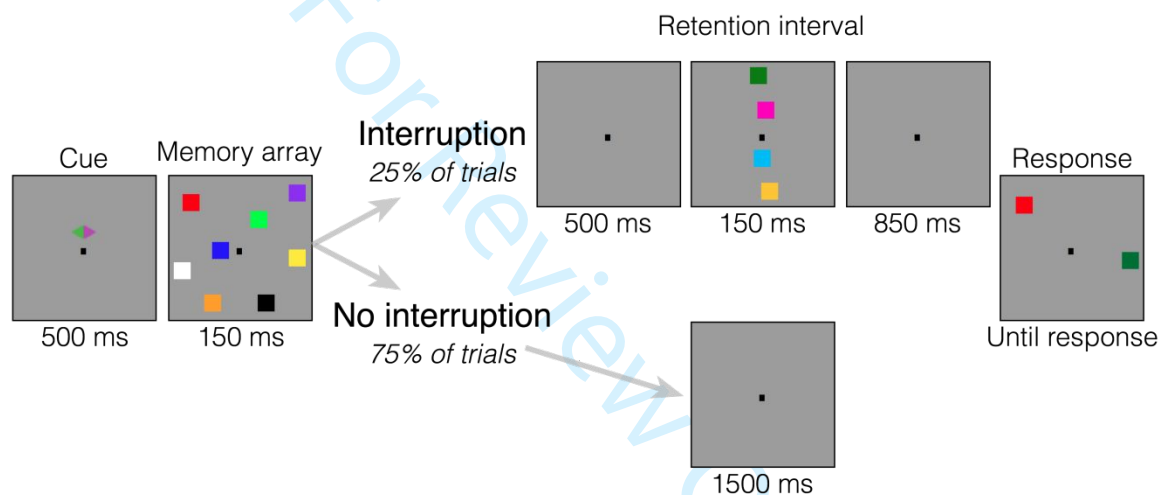


Figure 1. Task for Experiment 1. At the start of each trial, a cue appeared on the screen for 500 ms, which cued participants to attend one side of the screen. Then, an array of 4 colored squares briefly appeared (150 ms). On 75% of trials (no interruption condition), the retention interval (1500 ms) remained blank the entire time. On the other 25% of trials (interruption condition), the retention interval was blank for 500 ms, but then a series of four colored squares appeared on the midline for 150 ms. Participants were told to always ignore these squares that appeared on the midline of the screen during the retention interval. Following the brief interruption, the screen then went blank again for 850 ms. At the end of each trial, a response screen appeared with one square in each hemifield. Participants were told to report whether the square on the attended side was the same color as the original square in that location.

Behavioral Data Analysis.

We separately analyzed performance for the trials with and without interruption. Performance was converted to a capacity score, K , calculated as $N \times (H - FA)$, where N is the set-size, H is the hit rate, and FA is the false alarm rate (Cowan, 2011). To compare performance between the two conditions, we used a two-tailed, repeated measures t-test.

INTERRUPTION OF ONLINE REPRESENTATIONS

8

Artifact Rejection.

We recorded EEG activity from 30 active Ag/AgCl electrodes (Brain Products actiCHamp, Munich, Germany) mounted in an elastic cap positioned according to the International 10-20 system [Fp1, Fp2, F7, F8, F3, F4, Fz, FC5, FC6, FC1, FC2, C3, C4, Cz, CP5, CP6, CP1, CP2, P7, P8, P3, P4, Pz, PO7, PO8, PO3, PO4, O1, O2, Oz]. FPz served as the ground electrode and all electrodes were referenced online to TP10 and re-referenced off-line to the average of all electrodes. Incoming data were filtered [low cut-off = .01 Hz, high cut-off = 80 Hz, slope from low- to high-cutoff = 12 dB/octave] and recorded with a 500 Hz sampling rate. Impedances were kept below 10k Ω . To identify trials that were contaminated with eye movements and blinks, we used electrooculogram (EOG) activity and eye tracking. We collected EOG data with 5 passive Ag/AgCl electrodes (2 vertical EOG electrodes placed above and below the right eye, 2 horizontal EOG electrodes placed ~1 cm from the outer canthi, and 1 ground electrode placed on the left cheek). We collected eye-tracking data using a desk-mounted EyeLink 1000 Plus eye-tracking camera (SR Research Ltd., Ontario, Canada) sampling at 1,000 Hz. Usable eye-tracking data were acquired for 20 out of 22 participants in Experiment 1 and 29 out of 30 participants in Experiment 2.

EEG was segmented off-line with segments time-locked to memory display onset. Eye movements, blinks, blocking, drift, and muscle artifacts were first detected by applying automatic detection criteria to each segment. After automatic detection (see below), trials were manually inspected to confirm that detection thresholds were working as expected. Participants were excluded if they had less than 100 correct trials remaining in any of the conditions. For the participants used in analyses, we rejected on average 21% of trials in Experiment 1, and 39% of trials in Experiment 2.

Eye movements. We used a sliding window step-function to check for eye movements in the HEOG and the eye-tracking gaze coordinates. For HEOG rejection, we used a split-half sliding window approach. We slid a 100 ms time window in steps of 10 ms from the beginning to the end of the trial. If the change in voltage from the first half to the second half of the window was greater than 20 μ V, it was marked as an eye movement and rejected. For eye-tracking rejection, we applied a sliding window analysis to the x-gaze coordinates and y-gaze coordinates (window size = 100 ms, step size = 10 ms,

INTERRUPTION OF ONLINE REPRESENTATIONS

9

1
2
3
4 threshold = 0.5° of visual angle).

5 **Blinks.** We used a sliding window step function to check for blinks in the VEOG
6 (window size = 80 ms, step size = 10 ms, threshold = $30 \mu\text{V}$). We checked the eye-tracking
7 data for trial segments with missing data-points (no position data is recorded when the eye
8 is closed).
9

10
11
12 **Drift, muscle artifacts, and blocking.** We checked for drift (e.g. skin potentials)
13 by comparing the absolute change in voltage from the first quarter of the trial to the last
14 quarter of the trial. If the change in voltage exceeded $100 \mu\text{V}$, the trial was rejected for
15 drift. In addition to slow drift, we checked for sudden step-like changes in voltage with a
16 sliding window (window size = 100 ms, step size = 10 ms, threshold = $100 \mu\text{V}$). We
17 excluded trials for muscle artifacts if any electrode had peak-to-peak amplitude greater
18 than $200 \mu\text{V}$ within a 15 ms time window. We excluded trials for blocking if any electrode
19 had at least 30 time-points in any given 200-ms time window that were within 1V of each
20 other.
21
22
23
24
25
26
27

Contralateral Delay Activity Analysis

28
29 Segmented EEG data was baselined from 200 ms to 0 ms before the onset of the
30 memory displays. Artifact-free EEG segments from correct trials were averaged separately
31 for each condition (no interruption, interruption) and separately for electrodes ipsi- and
32 electrodes contralateral to the attended side. Then the difference between contra- and ipsi-
33 lateral activity for the electrode pair PO7/PO8 was calculated (i.e., the CDA), resulting in
34 two average waveforms for each participant. The average CDA amplitude was calculated
35 for three time windows: before interruption onset (450-650 ms), after interruption onset
36 (800-1,000 ms), before probe onset (1,300-1,500 ms). We then compared the CDA for each
37 time window with a repeated measures two-tailed t-test. To measure the robustness of the
38 CDA for each condition (reliable difference between contra- and ipsilateral activity), we
39 also ran a one-sample t-test (against zero) for each time window.
40
41
42
43
44
45
46
47

Lateralized Alpha Power Analysis

48
49 The same EEG segments as the CDA analysis were used in this analysis, however,
50 the segments were not baselined. The raw EEG signal was band-pass filtered in the alpha
51 band (8-12 Hz) using a two-way least-squares finite-impulse-response filter (“eegfilt.m”
52 from EEGLAB Toolbox). Instantaneous power was then extracted by applying a Hilbert
53
54
55
56
57
58
59
60

INTERRUPTION OF ONLINE REPRESENTATIONS

10

transform ('hilbert.m') to the filtered data. The resulting data were averaged separately for each condition (no interruption, interruption) and each laterality (contra- versus ipsi-lateral to cued hemifield) for the electrode pair PO7/PO8, resulting in four average waveforms for each participant. The average alpha power was calculated for the same three time windows as the CDA analysis. We then compared lateralized alpha power suppression for each time window with a repeated-measures two-tailed t-test. To measure the robustness of lateralized alpha power suppression for each condition (reliable difference between contra- and ipsilateral activity), we also ran a one-sample t-test (against zero) for each time window.

Results

Behavior

Performance (Figure 2), as measured by K, was significantly worse on trials that were interrupted ($M=1.6$), than on trials that were not interrupted ($M=1.2$), significant two-way repeated measures t-test, $t(19)=4.428$, $p<0.001$, $M=0.408$, 95% CI (0.215, 0.601).

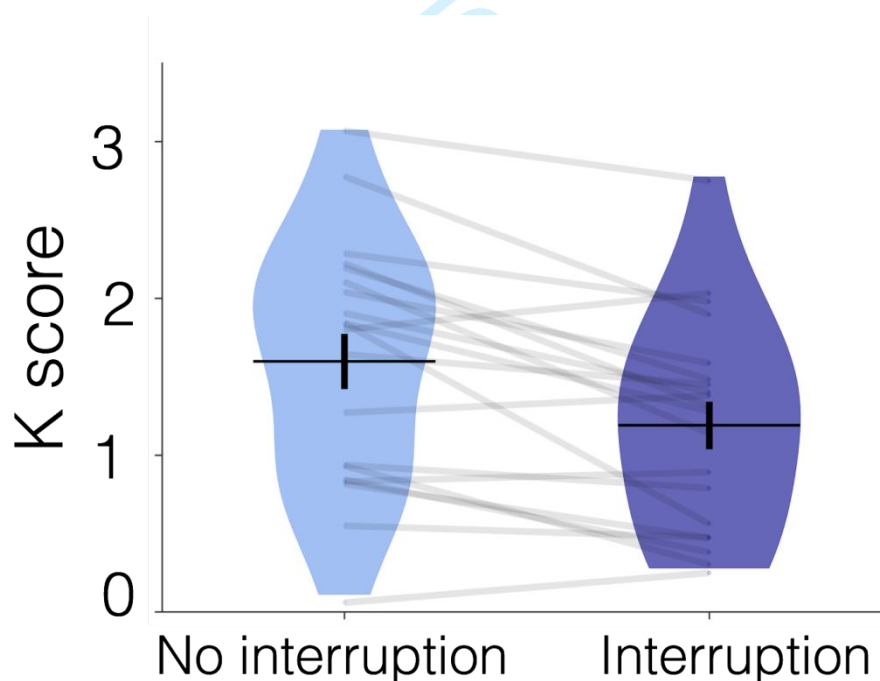


Figure 2. Behavioral performance for Experiment 1. Average performance is represented with the horizontal black line. The standard error is represented as the vertical black line. The distribution of K scores for all participants is represented by the violin plot. Dots and light gray lines represent one participant's performance.

Contralateral Delay Activity

INTERRUPTION OF ONLINE REPRESENTATIONS

11

1
2
3
4 **Pre-Interruption (450-650 ms).** The CDA (Figure 3) was robust before
5 interruption onset (450-650 ms) on trials with, $t(19)=-3.187$, $p=0.005$, $M=-0.707$, 95% CI
6 $(-1.171 -0.243)$, and without, $t(19)=-4.053$, $p=0.001$, $M=-0.837$, 95% CI $(-1.270 -0.405)$
7 interruption. There was not a significant difference in CDA amplitude between trials with
8 and without interruption during this time window, $t(19)=-1.394$, $p=0.179$, $M=-0.131$, 95%
9 CI $(-0.327, 0.066)$.

10
11
12
13
14 **Post-Interruption I (800-1,000 ms).** Immediately following the offset of the
15 interruption (800-1,000 ms), the CDA remained robust for both conditions, No
16 Interruption: $t(19)=-4.016$, $p=0.001$, $M= -0.674$, 95% CI $(-1.025 -0.323)$; Interruption:
17 $t(19)=-2.928$, $p=0.009$, $M=-0.583$, 95% CI $(-1.000 -0.166)$. Again, there was not a
18 significant difference in CDA amplitude between trials with and without interruption,
19 $t(19)= -0.525$, $p= 0.606$, $M=-0.091$, 95% CI $(-0.452 0.271)$.

20
21
22
23 **Post-interruption II (1,300-1,500 ms).** By the end of the trial (1,300-1,500 ms),
24 however, there was a significant difference in CDA amplitude between trials with and
25 without interruption, $t(19)= -5.145$, $p= <0.001$, $M=-0.731$, 95% CI $(-1.028, -0.434)$. On
26 trials without interruption, the CDA remained robust, $t(19)=-2.535$, $p=0.020$, $M= -0.496$,
27 95% CI $(-0.906 -0.086)$. However, on trials with interruption, the CDA was no longer
28 significantly different from zero, $t(19)= 1.445$, $p=0.165$, $M=0.234$, 95% CI $(-0.105 0.574)$.
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

INTERRUPTION OF ONLINE REPRESENTATIONS

12

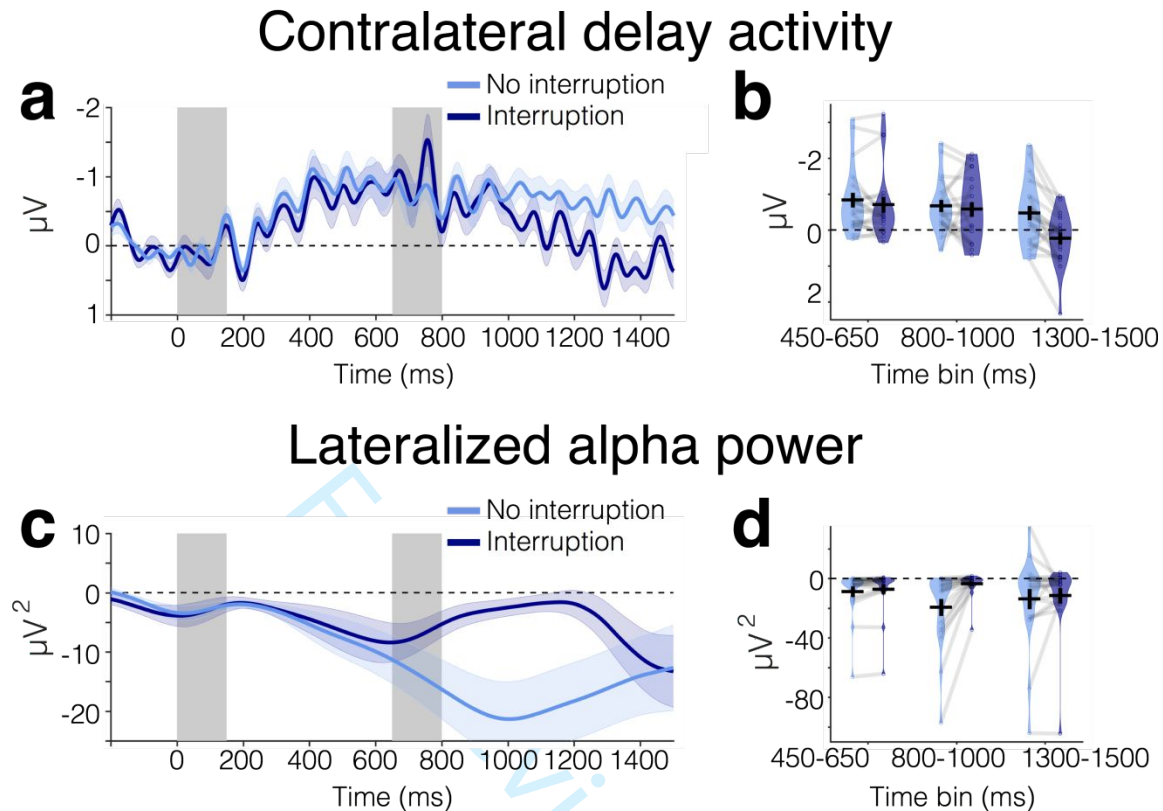


Figure 3. Contralateral delay activity and lateralized alpha power for Experiment 1. (a) CDA amplitude over time for trials with (dark blue line) and without (light blue line) interrupters. The light color envelopes around each line represent standard error of the mean for each condition. The first vertical gray bar (timepoint 0-150 ms) represents when the memory array was on the screen, and the second gray bar (timepoint 650-800 ms) represents when the interrupters were on the screen, if there were interrupters on that trial. (b) CDA amplitude for trials with (light blue plots) and without interruption (dark blue plots) averaged over the three time windows of interest (450-650 ms, 800-1000 ms, 1300-1500 ms). Each dot represents the mean number correct in one condition for one participant. A line connects data from the same participant. The horizontal black line depicts the mean across participants. The black error bars reflect the standard error of the mean. The shaded area reflects the density across participants. (c) Lateralized alpha power over time for trials with and without interruption (d) Average lateralized alpha power for trials with and without interruption.

Lateralized Alpha Power

Pre-Interruption (450-650 ms). Alpha power (Figure 3) was significantly more negative at contralateral compared to ipsilateral electrodes before interruption onset (450-650 ms) on trials with, $t(19)=-2.131$, $p=0.046$, $M=-7.264$, 95% CI (-14.398 -0.130), and without interruption, $t(19)=-2.517$, $p=0.021$, $M=-8.815$, 95% CI (-16.145 -1.486). Alpha power was significantly more lateralized on trials without interruption than trials with interruption during this time window, $t(19)=-2.573$, $p=0.019$, $M=-1.551$, 95% CI (-2.813

INTERRUPTION OF ONLINE REPRESENTATIONS

13

-0.289). We suspect that this may be due to time smearing in the alpha analysis. Time smearing is a side-effect of Fourier transforms, as the calculation of power at any time point incorporates data from time points before and after the time point of interest. Therefore, the effect of the interruption may be smeared across time, causing it to appear like there are differences in alpha power before interruption onset when there actually are only differences after interruption onset.

Post-Interruption I (800-1,000 ms). Immediately following the offset of the interruption (800-1,000 ms), lateralization of alpha power remained robust following trials without interruption, $t(19)=-3.423$, $p=0.003$, $M=-19.393$, 95% CI (-31.250 -7.536). However, alpha power was not significantly lateralized following trials with interruption, but this effect was trending towards significance, $t(19)=-2.054$, $p=0.054$, $M=-3.554$, 95% CI (-7.175 0.067). During this time window, lateralized alpha power was significantly more lateralized on trials without interruption than trials with interruption, $t(19)=-3.629$, $p=0.002$, $M=-15.839$, 95% CI (-24.974 -6.704).

Post-interruption II (1,300-1,500 ms). By the end of the trial (1,300-1,500 ms), however, there was no longer a significant difference in lateralized alpha power between the two conditions, $t(19)=-0.904$, $p=0.377$, $M=-3.879$, 95% CI (-12.863 5.104). Lateralized alpha power was significantly lateralized in both conditions, No Interruption: $t(19)=-2.144$, $p=0.045$, $M=-14.207$, 95% CI (-28.077 -0.337), Interruption: $t(19)=-2.137$, $p=0.046$, $M=-10.328$, 95% CI (-20.444 -0.213).

Conclusions

In Experiment 1, participants' working memory performance was reduced when they were interrupted during the retention interval. In addition, both the CDA and lateralized alpha power were negatively impacted by the interrupters, but this effect had distinct time courses for the two signals. The CDA briefly sustained following interruption while alpha power immediately became less lateralized. By the end of the trial, CDA was no longer present, but alpha power re-lateralized. These results suggest that the CDA and lateralized alpha power respond distinctly to task-irrelevant interruptions.

Experiment 2

INTERRUPTION OF ONLINE REPRESENTATIONS

14

1
2
3
4 In Experiment 2, we sought to determine whether the neural responses to task-
5 irrelevant interruptions could be modulated by task expectancy, or if they are fixed
6 responses to interruption irrespective of the subject's expectations. Thus, in Experiment 2,
7 we compared the same 25% interruption condition employed in Experiment 1 with one in
8 which interrupters were presented on 75% of trials. We predicted that a higher frequency
9 of task-irrelevant interruptions should allow participants to be better prepared for
10 interruptions. Accordingly, CDA and lateralized alpha power should sustain for longer
11 following interruption. Importantly, we will also examine the onset and offset of the CDA
12 and lateralized alpha power to examine whether the time course of the two subprocesses
13 may be affected differently.
14
15
16
17
18
19
20

Materials & Methods

21
22
23 **Participants.** Thirty novel volunteers, naïve to the objective of the experiment
24 participated for payment (~15 USD per hour). All data were collected in a single session.
25 The data of 10 participants were excluded from the analysis because of too many artifacts,
26 poor behavioral performance or technical problems (same criteria as in Experiment 1). The
27 remaining 20 participants (11 male) were between the ages of 19-32 ($M = 23.54$, $SD =$
28 3.85).
29
30
31
32

33 **Stimuli & Procedures.** Stimuli were identical to Experiment 1 (Figure 1).
34 Procedure was also identical to Experiment 1, except for the following changes. The
35 retention interval was increased to 2000 ms. The experiment was divided in two halves in
36 each of which the probability of interruption was varied. The order of the halves was
37 counterbalanced across participants. The probability for interruption was 25% in one part
38 (no interruption: 75%) and 75% in the other part (no interruption: 25%). This resulted in 2
39 \times 2 design with the factors Interruption (no interruption versus interruption) and Probability
40 (high versus low). Participants completed 1,920 trials in total (24 blocks of 80 trials each),
41 240 trials for each of the two low probability conditions and 720 trials for each of the two
42 high probability conditions.
43
44
45
46
47
48
49

Analysis

50
51
52 Behavioral and EEG data were analyzed analogously to Experiment 1, but included
53 the additional factor Probability. For statistical analyses, we forwarded the mean CDA
54 amplitude (contra- minus ipsilateral activity) to a two-way ANOVA with the within-
55
56
57
58
59
60

INTERRUPTION OF ONLINE REPRESENTATIONS

15

subjects factors Interruption (interruption versus no interruption) and Probability (75% probability for interruption versus 25%). Additionally, for the CDA and lateralized alpha analyses, the time window before probe onset was 1800-2000 ms, as we extended the retention interval by 500 ms.

Results**Behavior**

Performance (Figure 4), as measured by K, was significantly worse on trials that were interrupted (Low Probability $M = 1.4$, High Probability $M = 1.6$) than on trials that were not interrupted (Low Probability $M = 1.6$, High Probability $M = 1.7$), regardless of probability, significant main effect of Interruption, $F(1,19)=21.288$, $p<0.001$, $\eta_p^2=0.528$. There was not a significant main effect of Probability ($F(1,19)=3.575$, $p=0.074$, $\eta_p^2=0.158$) or interaction of Interruption and Probability ($F(1,19)=1.420$, $p=0.248$, $\eta_p^2=0.070$).

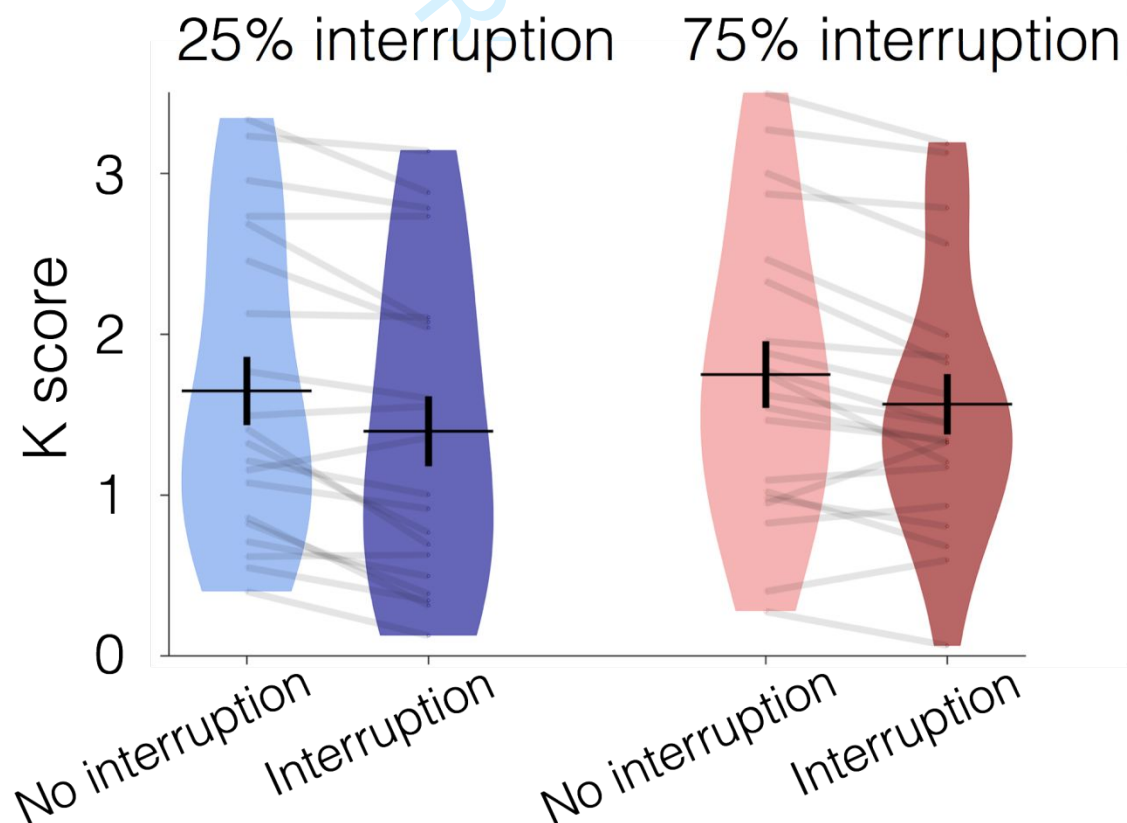


Figure 4. Behavioral performance for Experiment 2. Average performance is represented with the horizontal black line. The standard error is represented as the vertical black line. The distribution of K scores for all participants is represented by the violin plot. Dots and light gray lines represent one participant's performance.

INTERRUPTION OF ONLINE REPRESENTATIONS

16

Contralateral Delay Activity

Pre-interruption II (450-650 ms). Before interruption onset (450-650 ms), there was a significant CDA in all four conditions, all one-sample t-tests $p \leq 0.002$.

Additionally, there was no difference in CDA amplitude between any of the conditions, $p \geq 0.060$ for the main effects of Interruption, Probability and their interaction. Though the interaction of Interruption and Probability was trending towards significance, $p = 0.060$, $\eta_p^2 = 0.173$.

Post-interruption I (800-1,000 ms). Immediately following interruption offset (800-1,000 ms), the influence of interruption on CDA amplitude depended on the probability of being interrupted, significant interaction of Probability and Interruption: $F(1,19) = 9.951$, $p = 0.005$, $\eta_p^2 = 0.344$. Follow-up t-tests run separately for trials with and without interruption revealed that when interrupters were present, the amplitude of the CDA depended on the probability of interruption, $t(19) = 2.252$, $p = 0.036$. The CDA was significantly larger in the High Probability condition ($M = -0.660$) than in the Low Probability condition ($M = -0.265$). On trials without interruption, there was no difference in CDA amplitude between probabilities, $t(19) = -0.858$, $p = 0.402$. The main effects of Interruption and Probability were not significant, $p \geq 0.129$.

Post-interruption II (1,800-2,000 ms). By the end of the trial (1,800-2,000 ms), there was no difference in CDA amplitude between any of the conditions, $p \geq 0.142$ for the main effects of Interruption, Probability, and their interaction. There was no longer a significant CDA in any condition, all one-way t-tests $p \geq 0.088$. CDA tends to decline over time, and by extending the delay period compared to Experiment 1, we may have reached the point at which the CDA tends to decline naturally (Vogel and Machizawa, 2004).

INTERRUPTION OF ONLINE REPRESENTATIONS

17

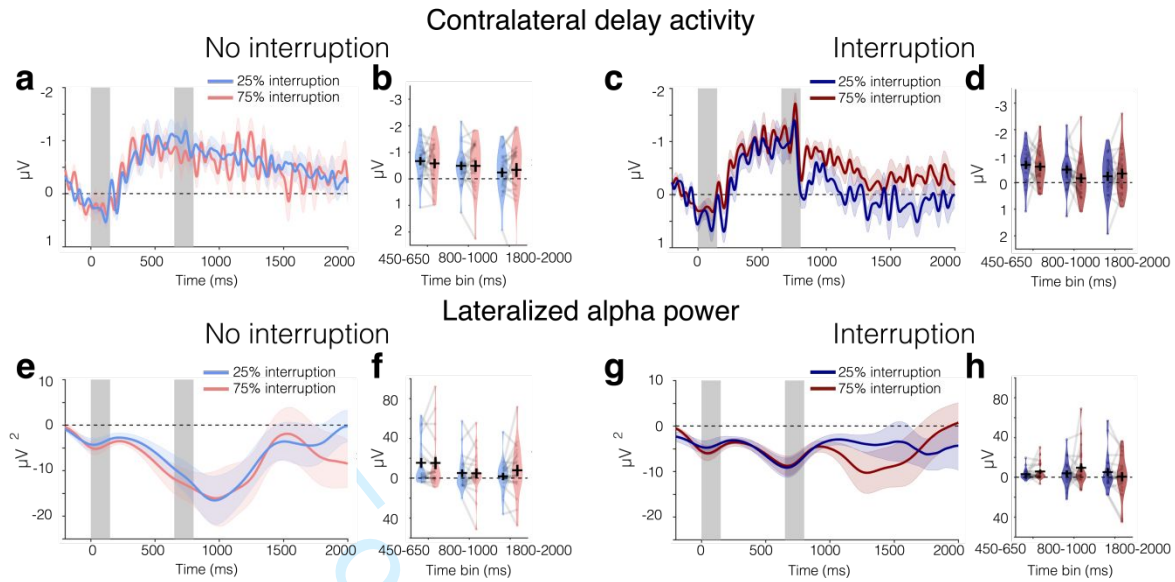


Figure 5. Contralateral delay activity and lateralized alpha power for Experiment 2. (a) CDA amplitude and (e) alpha power lateralization over time for trials without interruption in the 25% (light blue) and 75% (pink) interruption blocks. The light color envelopes around each line represent standard error of the mean for each condition. The first vertical gray bar (timepoints 0-150 ms) represents when the memory array was on the screen, and the second gray bar (timepoints 650-800 ms) represents when the interrupters were on the screen, if there were interrupters on that trial. (b) CDA amplitude and (f) alpha power lateralization for trials without interruption in the 25% (light blue plots) and 75% (pink plots) interruption blocks averaged over the three time windows of interest (450-650 ms, 800-1000 ms, 1800-2000 ms). Each dot represents the mean number correct in one condition for one participant. A line connects data from the same participant. The horizontal black line depicts the mean across participants. The black error bars reflect the standard error of the mean. The shaded area reflects the density across participants. (c) CDA amplitude and (g) lateralized alpha power over time for trials with interruption in the 25% (dark blue) and 75% (red) interruption blocks. (d) CDA amplitude and (h) alpha power lateralization averaged over three time windows of interest for trials with interruption in the 25% (dark blue plot) and 75% (red plot) interruption blocks.

Lateralized Alpha Power

Pre-interruption II (450-650 ms). Alpha power (Figure 5) was significantly suppressed in all conditions before the interruption onset (450-650 ms), all one-sample t -tests $p \leq 0.005$. The influence of interruption on alpha power suppression depended on the probability of interruption, significant interaction of Probability and Interruption: $F(1,19)=4.881$, $p=0.040$, $\eta_p^2 = 0.204$. As in Experiment 1, this pre-interruption difference could be due to time smearing of alpha power. There was no difference in lateralized alpha power between trials that were and were not interrupted for either the High ($f(19)=-2.046$,

INTERRUPTION OF ONLINE REPRESENTATIONS

18

p=0.055) or Low ($f(19)=0.279$, $p=0.789$) probability trials, though this effect was trending in the High probability trials.

Post-interruption I (800-1,000 ms). Immediately following the offset of interruption (800-1,000 ms), alpha power was significantly lateralized in all conditions, all one-sample t-tests $p \leq 0.006$. There was a significant main effect of Interruption on the strength of alpha lateralization: $F(1,19)=6.530$, $p=0.019$, $\eta_p^2 = 0.256$. For both High and Low probability trials, lateralized alpha power was stronger on trials without interruption (Low Probability $M = -15.498$, High Probability $M = -15.368$) than on trials with interruption (Low Probability $M = -4.549$, High Probability $M = -4.966$). No other effects were significant, $p \geq 0.778$.

Post-interruption II (1,800-2,000 ms). By the end of the trial (1,800-2,000 ms), the influence of interruption on lateralized alpha power depended on the probability of interruption, significant interaction of Interruption and Probability, $F(1,19)=6.365$, $p=0.021$, $\eta_p^2 = 0.251$. The difference between lateralized alpha power in high and low probability trials is different with and without interruption. However, within High probability trials, follow-up t-tests revealed that there was not an alpha power suppression difference between trials that were interrupted and those that were not for either High ($t(19)=-1.923$, $p=0.070$) or Low ($t(19)=1.077$, $p=0.295$) probability trials. Also, alpha power suppression was not significantly different between High and Low probability trials for trials that were interrupted ($t(19)=-1.278$, $p=0.217$) or for those that were not interrupted ($t(19)=1.362$, $p=0.189$). The interaction of Interruption and Probability is disordinal because the interaction is significant, but the follow-up t-tests are not significant. Disordinal interactions indicate that a factor has one kind of effect in one condition and the opposite kind of effect in the other condition. In this case, in the Low probability condition, alpha power is numerically more lateralized when interrupters were present ($M=-5.14$, $SD=18.89$) than when they were not present ($M=-1.84$, $SD=16.53$). However, in the High probability condition, alpha power is numerically less lateralized when interrupters were present ($M=-0.58$, $SD=20.83$) than when they were not present ($M=-7.296$, $SD=25.92$). The main effects of Interruption and Probability were also not significant, $p \geq 0.472$.

Conclusions

INTERRUPTION OF ONLINE REPRESENTATIONS

19

1
2
3
4 In Experiment 2, we found that behavioral performance was worse when
5 participants were interrupted than when they were not interrupted regardless of the
6 probability of interruption. Once again, the neural results revealed that both CDA and
7 lateralized alpha power were negatively impacted by interruptions, but these two signals
8 had distinct time courses. Following interruption, CDA sustained, but lateralized alpha
9 power became less lateralized. Additionally, the amplitude of CDA immediately following
10 interruption depended on the probability of interruption – CDA amplitude was larger when
11 participants were expecting to be interrupted. However, alpha power lateralization did not
12 depend on expectations – alpha power shifted towards baseline when interruptions were
13 present, regardless of the probability of interruption.
14
15
16
17
18
19
20

21 By the end of the trial, CDA was no longer present on trials with interruptions,
22 regardless of probability. This replicates the CDA results from Experiment 1. The effect of
23 probability and interruption on alpha power lateralization by the end of the trial were a bit
24 more ambiguous. In the Low probability block, lateralized alpha power was equivalent on
25 trials with and without interruptions. This replicates the results from Experiment 1.
26 However, upon visual inspection of the results, the “recovery” pattern following
27 interruption was not as apparent. This is because overall alpha power lateralization on trials
28 without interruption was very close to baseline, unlike in Experiment 1 where alpha power
29 was robustly lateralized. This reduction in alpha power lateralization on trials without
30 interruption could plausibly be due to the length of the retention interval. In Experiment 1,
31 the retention interval was 1,500 ms, and in Experiment 2 it was extended to 2,000 ms. We
32 did this so that we could investigate whether CDA would return if participants had more
33 time post-interruption. However, both CDA and alpha power lateralization tend to shift
34 towards baseline with longer delays, which may be the reason why alpha power is less
35 lateralized by the end of the trial in Experiment 2 than in Experiment 1. Regardless of the
36 amount of alpha power lateralization at the end of the trial, we still found a significant
37 reversal of the effect of probability and interruption in the High probability condition as
38 compared to the Low probability condition. In the High probability condition, alpha power
39 was more lateralized on trials without interruptions than on trials with interruptions, but
40 this effect was reversed in the Low probability condition.
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

INTERRUPTION OF ONLINE REPRESENTATIONS

20

General Discussion

Working memory maintains information so that it can be used despite momentary perturbations from task-irrelevant information. Here, we examined how memory representations that have already reached a stable state respond to visual interruption. As expected, we found a modest behavioral impact of interruption. Participants remembered significantly fewer items when they were interrupted than when they were not interrupted, but they performed above chance in all conditions. Despite a modest behavioral impact, task-irrelevant interruption produced substantial perturbations on two well-characterized EEG signals of working memory, lateralized alpha power and contralateral delay activity (CDA). Both lateralized alpha power, an index of sustained spatial attention, and CDA, an index of actively maintained working memory representations, were disrupted at certain points during the delay, but the time course of these perturbations varied. Lateralized alpha power results suggest that attention shifted towards baseline immediately following the interruption but had returned to the target positions by the end of the trial. By contrast, the CDA results suggest that working memory representations continued to persist following the interruption but was eliminated by the end of the trial. We additionally found that task expectancy modulated the timing and magnitude of these perturbations of working memory representations, suggesting that the brain's response to task-irrelevant interruption is regulated by task context. The distinct time courses of and the influence of task context on lateralized alpha power and CDA have many interesting theoretical implications that future work can help elucidate.

Neural response immediately following interruption

Sudden onsets of task-irrelevant interruption have been shown to capture attention when interrupters are visually salient (Bisley and Goldberg, 2003; Andrews, Ratwani and Trafton, 2009; van Moorselaar *et al.*, 2018). In our experiment, we used lateralized alpha power as an index of sustained spatial attention (Foster *et al.*, 2016; Hakim *et al.*, 2018). Following the onset of interruption, lateralized alpha power almost immediately shifted towards baseline. When lateralized alpha power is at baseline, it suggests that participants are no longer spatially attending the lateral memory items. Neural evidence from previous studies suggests that participants attend to the locations of interrupting stimuli (Bisley and

INTERRUPTION OF ONLINE REPRESENTATIONS

21

Goldberg, 2003; van Moorselaar *et al.*, 2018) because of attentional capture (Sawaki and Luck, 2012; Feldmann-Wüstefeld and Schubö, 2013). Thus, in the present study, participants presumably shifted their attention away from lateralized representations following the onset of task-irrelevant interruption to the centrally presented interrupters.

During this same time window, CDA remained robust and significantly above baseline. Previous research has shown that CDA is sensitive to trial-by-trial fluctuations in working memory performance and tracks the number of maintained object representations (Ikkai, McCollough and Vogel, 2010; Adam *et al.*, 2015). Considering this, the robust CDA immediately following the onset of interruption suggests that object representations persist, at least momentarily, following the withdrawal of spatial attention to a new position. The presence of CDA and lack of lateralized alpha power immediately following interruption raises the long-standing theoretical question of whether object representations maintained in working memory can persist without sustained spatial attention. Previous research has suggested that spatial attention is a rehearsal mechanism that facilitates the maintenance of object representations held in working memory (Williams and Woodman, 2012). Additionally, the positions of object representations are maintained in working memory even when spatial information is completely irrelevant (Foster *et al.*, 2017). Together, these previous results suggest that spatial attention aids the maintenance of working memory information, but do not address whether working memory representations necessitate sustained spatial attention. In the present study, the robust CDA and lack of lateralized alpha power following the onset of interruption suggest that object representations maintained in working memory can persist without sustained spatial attention. Therefore, our results suggest that working memory representations may not necessitate sustained spatial attention. Nevertheless, working memory representations may still be volatile without sustained spatial attention, given that CDA goes to baseline by the end of trials with interruption.

Neural activity at the end of interrupted trials

In this study, we sought to interrupt participants after working memories reached a stable state. Therefore, it is not surprising that participants can still perform well above chance in the distractor-present trials. It is likely that interruptions to the WM

INTERRUPTION OF ONLINE REPRESENTATIONS

22

1
2
3
4 representations at earlier moments, such as before CDA is fully formed, would produce
5 larger behavioral decrements (e.g., Vogel et al 2006). Nevertheless, by the end of
6 interrupted trials, we observed that the CDA was no longer reliable, but alpha power
7 became re-lateralized. There is a large body of research that has shown that CDA tracks
8 the active maintenance of information, is sensitive to trial-by-trial fluctuations in working
9 memory performance, and distinguishes stable individual differences in working memory
10 (Vogel and Machizawa, 2004; Luria *et al.*, 2016). Therefore, the pattern of activity at the
11 end of the trial suggests that participants re-oriented their attention to the locations of the
12 memoranda, but no longer maintained active working memory representations. If
13 participants no longer maintained object representations that are tracked by CDA, how
14 were they able to still perform the change detection task on interrupted trials (albeit worse
15 than non-interrupted trials)? There are a few possible explanations.

24 One possible explanation for the absence of the CDA at the end of the trial but
25 above chance behavioral performance is that performance on interrupted trials could rely
26 on offline memory representations. Previous research has shown that information in
27 working memory can be simultaneously maintained in both active and passive memory
28 states (Mallett & Lewis-Peacock, 2018). Therefore, when actively maintained memory
29 traces are no longer present, information could still be retrieved from an offline state.
30 Research that has investigated retrieval of information from offline memory states has
31 found that alpha power tracks information retrieved from long-term memory (Fukuda,
32 Kang and Woodman, 2016). Additionally, other research has suggested that attention can
33 aid recall of information that would be otherwise unavailable to working memory (Murray
34 et al, 2013). These findings dovetail with our results – at the end of interrupted trials, when
35 information about the memoranda is required to respond to the probe, lateralized alpha
36 power could be re-instated in order to re-load information from offline memory storage,
37 thereby bolstering behavioral performance. An alternate explanation for the recovery of
38 lateralized alpha power at the end of the trial is that it reflects the anticipation of the
39 upcoming memory probe. The memory probe always appeared in the same location as one
40 of the memory items. Thus, in order to shift attention to the location of the upcoming probe,
41 participants had to remember the locations of the original memory items. Therefore, even
42 if the re-lateralization of alpha power at the end of interrupted trials reflects the orienting
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

INTERRUPTION OF ONLINE REPRESENTATIONS

23

of spatial attention to the location of the anticipated memory probe, it still suggests that this re-lateralization relies on the retrieval of task-relevant spatial information. Both the re-loading and re-orienting explanations of the recovery of alpha power are plausible and theoretically interesting explanations for this pattern of activity.

The relatively good behavioral performance without CDA could alternately be explained by other neural traces of actively maintained working memory representations that we are not measuring. The CDA is a coarse neural measure that compares activity contralateral and ipsilateral to memory items. Thus, it is not an exhaustive measure of working memory. More spatially global neural signals or more distributed patterns of activity, for example, could sustain following task-irrelevant interruption, and these signals could plausibly bolster behavioral performance. Regardless of the mechanism that preserves information about the memoranda, our results strongly suggest that actively maintained information is dynamically perturbed following task-irrelevant interruption.

Modulation of CDA and alpha power by task demands

In Experiment 2, we varied task demands by interrupting participants on 75% (high) or 25% (low) of trials. Following interruption onset, we found the same pattern of result as Experiment 1; Lateralized alpha power shifted towards baseline while CDA persisted. However, the amplitude of the CDA varied as a function of task demands. When task demands were high, CDA amplitude was higher than when task demands were low. This suggests that participants were able to better protect working memory representations when they were expecting to be interrupted, and that task context is involved in how the brain responds to task-irrelevant interruption. On the other hand, the influence of task demands on lateralization of alpha power was more ambiguous. Our results suggest that spatial attention may be uniformly captured by interrupters initially regardless of expectation. However, during certain points in the trial, lateralization of alpha power may vary as a function of task demands. Therefore, the neural responses to interruption that we observed were affected both by both interruption and task demands. These results go hand-in-hand with previous research that has shown that distraction by salient irrelevant stimuli can be modulated by top-down control. For example, when a color singleton is presented on 20% of the trials, it slows down response times in a visual search task more than when it is

INTERRUPTION OF ONLINE REPRESENTATIONS

24

presented on 50% of trials (Horstmann, 2005; Müller *et al.*, 2009; Marini, Chelazzi and Maravita, 2013; Folk and Remington, 2015) because attention requires more time to be deployed to the relevant information when rare distractors appear (Töllner, Müller and Zehetleitner, 2012).

Conclusions

In this set of experiments, we investigated the impact of task-irrelevant interruption on two dissociable neural signals, CDA, a neural index of actively maintained representations, and lateralized alpha power, an index of sustained spatial attention. By tracking these neural markers of working memory, we were able to observe changes in active representations that would not be apparent from behavioral measures alone. Both CDA and lateralized alpha power were impacted by task-irrelevant information, yet had distinct time courses. Our results suggest that following interruption, lateralized visual representations of memoranda can stay active in working memory for a short period of time without lateralized spatial attention before they are lost. These representations do not recover by the end of the trial, and are presumed to be stored offline. By contrast, attention is directed away from the spatial location of memoranda immediately after the onset of the interruption but can recover later and may even contribute to the retrieval of information from offline storage. Thus, our results show that task-irrelevant interruption could motivate the transfer of information from active to passive storage. Moreover, the dissociation between CDA and lateralized alpha power further emphasizes that these neural markers distinctly contribute to the maintenance of information in working memory and may distinctly protect actively maintained memories from interruption.

Bibliography

- Adam, K. C. S. *et al.* (2015) 'The contribution of attentional lapses to individual differences in visual working memory capacity', *Journal of Cognitive Neuroscience*. doi: 10.1162/jocn.
- Adam, K. C. S., Robison, M. K. and Vogel, E. K. (2018) 'Contralateral Delay Activity Tracks Fluctuations in Working Memory Performance Kirsten', *Journal of Cognitive Neuroscience*, 26(3), pp. 194–198. doi: 10.1162/jocn.
- Andrews, A., Ratwani, R. and Trafton, G. (2009) 'Recovering From Interruptions : Does Alert Type Matter ?', pp. 409–413.
- Bisley, J. W. *et al.* (2004) 'Activity of neurons in cortical area MT during a memory for motion task', *J. Neurophysiol.* Macmillan Publishers Limited. All rights reserved, 91, pp. 286–300. Available at: <http://dx.doi.org/10.1152/jn.00870.2003>.
- Bisley, J. W. and Goldberg, M. E. (2003) 'Neuronal Activity in the Lateral Intraparietal Area and Spatial Attention', 299(January).
- Chun, M. M. and Turk-Browne, N. B. (2007) 'Interactions between attention and memory', *Current Opinion in Neurobiology*, 17(2), pp. 177–184. doi: 10.1016/j.conb.2007.03.005.
- Clapp, W. C., Rubens, M. T. and Gazzaley, A. (2010) 'Mechanisms of working memory disruption by external interference', *Cerebral Cortex*, 20(4), pp. 859–872. doi: 10.1093/cercor/bhp150.
- Feldmann-Wüstefeld, T. and Schubö, A. (2013) 'Context homogeneity facilitates both distractor inhibition and target enhancement', *Journal of Vision*, 13(2013), pp. 1–12. doi: 10.1167/13.3.11.doi.
- Feldmann-Wüstefeld, T. and Vogel, E. K. (2018) 'Neural Evidence for the Contribution of Active Suppression During Working Memory Filtering', *Cerebral Cortex*. doi: 10.1093/cercor/bhx336.
- Feldmann-Wüstefeld, T., Vogel, E. K. and Awh, E. (2018) 'Contralateral Delay Activity Indexes Working Memory Storage, Not the Current Focus of Spatial Attention Tobias Feldmann-Wüstefeld, Edward K. Vogel, and Edward Awh Abstract', *Journal of Cognitive Neuroscience*2, pp. 1–11. doi: 10.1162/jocn.
- De Fockert, J. W. *et al.* (2001) 'The Role of Working Memory in Visual Selective Attention', *Science*, 291(January), pp. 1803–1806. doi: 10.1126/science.1056496.
- Folk, C. L. and Remington, R. W. (2015) 'Unexpected abrupt onsets can override a top-down set for color', *Journal of Experimental Psychology: Human Perception and Performance*, 41(4), pp. 1153–1165. doi: 10.1037/xhp0000084.
- Foster, J. J. *et al.* (2016) 'The topography of alpha-band activity tracks the content of spatial working memory.', *Journal of neurophysiology*, 115(1), pp. 168–77. doi: 10.1152/jn.00860.2015.
- Foster, J. J. *et al.* (2017) 'Alpha-Band Activity Reveals Spontaneous Representations of Spatial Position in Visual Working Memory', *Current Biology*. Elsevier Ltd., 27(20), p. 3216–3223.e6. doi: 10.1016/j.cub.2017.09.031.
- Fukuda, K., Kang, M.-S. and Woodman, G. F. (2016) 'Distinct neural mechanisms for spatially lateralized and spatially global visual working memory representations', *Journal of Neurophysiology*, 116(4), pp. 1715–1727. doi: 10.1152/jn.00991.2015.
- Fukuda, K., Mance, I. and Vogel, E. K. (2015) 'Alpha Power Modulation and Event-

INTERRUPTION OF ONLINE REPRESENTATIONS

26

1
2
3
4 Related Slow Wave Provide Dissociable Correlates of Visual Working Memory', *Journal*
5 *of Neuroscience*, 35(41), pp. 14009–14016. doi: 10.1523/JNEUROSCI.5003-14.2015.

6 Gaspar, J. M. and McDonald, J. J. (2014) 'Suppression of salient objects prevents
7 distraction in visual search', *The Journal of Neuroscience*, 34(16), pp. 5658–66. doi:
8 10.1523/JNEUROSCI.4161-13.2014.

9 Hakim, N. *et al.* (2018) 'Dissecting the neural focus of attention reveals distinct processes
10 for spatial attention and object-based storage in visual working memory.', *BioRxiv*.

11 Horstmann, G. (2005) 'Attentional capture by an unannounced color singleton depends
12 on expectation discrepancy', *Journal of Experimental Psychology: Human Perception*
13 *and Performance*, 31(5), pp. 1039–1060. doi: 10.1037/0096-1523.31.5.1039.

14 Ikkai, A., McCollough, A. W. and Vogel, E. K. (2010) 'Contralateral delay activity
15 provides a neural measure of the number of representations in visual working memory.',
16 *Journal of neurophysiology*, 103(4), pp. 1963–1968. doi: 10.1152/jn.00978.2009.

17 Jolicœur, P. and Dell'Acqua, R. (1998) 'The Demonstration of Short-Term
18 Consolidation', *Cognitive Psychology*, 36(2), pp. 138–202. doi: 10.1006/cogp.1998.0684.

19 Luria, R. *et al.* (2016) 'The contralateral delay activity as a neural measure of visual
20 working memory', *Neuroscience and Biobehavioral Reviews*, 34(5), pp. 352–359. doi:
21 10.1177/0963721414541462. Self-Control.

22 Marini, F., Chelazzi, L. and Maravita, A. (2013) 'The costly filtering of potential
23 distraction: Evidence for a supramodal mechanism', *Journal of Experimental*
24 *Psychology: General*, 142(3), pp. 906–922. doi: 10.1037/a0029905.

25 van Moorselaar, D. *et al.* (2017) 'Spatially Selective Alpha Oscillations Reveal Moment-
26 by-Moment Trade-offs between Working Memory and Attention', *Journal of Cognitive*
27 *Neuroscience*, 30(2), pp. 256–266. doi: 10.1162/jocn.

28 van Moorselaar, D. *et al.* (2018) 'Spatially Selective Alpha Oscillations Reveal Moment-
29 by-Moment Trade-offs between Working Memory and Attention', *Journal of Cognitive*
30 *Neuroscience*, 26(3), pp. 194–198. doi: 10.1162/jocn.

31 Müller, H. J. *et al.* (2009) 'Attentional Capture by Salient Color Singleton Distractors Is
32 Modulated by Top-Down Dimensional Set', *Journal of Experimental Psychology:*
33 *Human Perception and Performance*, 35(1), pp. 1–16. doi: 10.1037/0096-1523.35.1.1.

34 Olivers, C. N. L. (2008) 'Interactions between visual working memory and visual
35 attention', *Frontiers in Biological Science*, pp. 1182–1191.

36 Postle, B. R., D'Esposito, M. and Corkin, S. (2005) 'Effects of verbal and nonverbal
37 interference on spatial and object visual working memory', *Memory and Cognition*,
38 33(2), pp. 203–212. doi: 10.3758/BF03195309.

39 Sawaki, R. and Luck, S. J. (2012) 'Active suppression of distractors that match the
40 contents of visual working memory Risa', *Visual Cognition*, 19(7), pp. 1–14. doi:
41 10.1080/13506285.2011.603709. Active.

42 Thut, G. *et al.* (2006) 'Alpha Band Electroencephalographic Activity over Occipital
43 Cortex Indexes Visuospatial Attention Bias and Predicts Visual Target Detection',
44 *Journal of Neuroscience*, 26(37), pp. 9494–9502. doi: 10.1523/JNEUROSCI.0875-
45 06.2006.

46 Töllner, T., Müller, H. J. and Zehetleitner, M. (2012) 'Top-down dimensional weight set
47 determines the capture of visual attention: Evidence from the PCN component', *Cerebral*
48 *Cortex*, 22(7), pp. 1554–1563. doi: 10.1093/cercor/bhr231.

49 Vogel, E. K. and Machizawa, M. G. (2004) 'Neural activity predicts individual
50
51
52
53
54
55
56
57
58
59
60

INTERRUPTION OF ONLINE REPRESENTATIONS

27

1
2
3
4 differences in visual working memory capacity', *Nature*. Macmillan Publishers Limited.
5 All rights reserved, 428, pp. 748–751. Available at:
6 <http://dx.doi.org/10.1038/nature02447>.

7 Vogel, E. K., McCollough, A. W. and Machizawa, M. G. (2005) 'Neural measures reveal
8 individual differences in controlling access to working memory', *Nature*, 438(7067), pp.
9 500–503. doi: 10.1038/nature04171.

10 Vogel, E. K., Woodman, G. F. and Luck, S. J. (2006) 'The time course of consolidation
11 in visual working memory', *Journal of Experimental Psychology: Human Perception and*
12 *Performance*, 32(6), pp. 1436–1451. doi: 10.1037/0096-1523.32.6.1436.

13 Williams, M. and Woodman, G. F. (2012) 'Directed Forgetting and Directed
14 Remembering in Visual Working Memory', *Journal Of Experimental Psychology-*
15 *Learning Memory And Cognition*. doi: 10.1037/a0027389.

16 Woodman, G. F. and Chun, M. M. (2006) 'The role of working memory and long-term
17 memory in visual search', *Visual Cognition*, 14(4–8), pp. 808–830. doi:
18 10.1080/13506280500197397.

19 Worden, M. S. *et al.* (2000) 'Anticipatory biasing of visuospatial attention indexed by
20 retinotopically specific alpha-band electroencephalography increases over occipital
21 cortex.', *The Journal of neuroscience*, 20, pp. 1–6. doi: Rc63.

22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Review Only