Electrophysiological Correlates of Visual Memory Search

Lauren H. Williams[1]*, Iris Wiegand[2]*, Mark Lavelle[3], Jeremy M. Wolfe[4], Keisuke Fukuda[5], Marius V. Peelen[2], and Trafton Drew[6]

Abstract

■ In everyday life, we frequently engage in 'hybrid' visual and memory search, where we look for multiple items stored in memory (e.g., a mental shopping list) in our visual environment. Across three experiments, we used event-related potentials to better understand the contributions of visual working memory (VWM) and long-term memory (LTM) during the memory search component of hybrid search. Experiments 1 and 2 demonstrated that the FN400 – an index of LTM recognition – and the CDA –an index of VWM load – increased with memory set size (target load), suggesting that both VWM and LTM are involved in memory search, even when target load exceeds capacity limitations of VWM. In Experiment 3, we used these electrophysiological indices to test how categorical similarity of targets and distractors affects memory search. The CDA and FN400 were modulated by memory set size only if items resembled targets. This suggests that dissimilar distractor items can be rejected before eliciting a memory search. Together, our findings demonstrate the interplay of VWM and LTM processes during memory search for multiple targets. ■

MEMORY IN MULTIPLE-TARGET SEARCH

When shopping at the grocery store, you likely have a shopping list of multiple items (e.g., noodles, spaghetti sauce, and bread) you intend to purchase. This routine task involves a "hybrid" search through memory (i.e., your mental shopping list) and through the visual environment (i.e., the items at the grocery store; Wolfe, 2012, 2021. The memory search component of this task involves recognition of one of many target items, retrieved from memory, while attending to items in the visual space. Hybrid search is also an important task for many professionals, such as radiologists and baggage screeners, who need to keep an eye out for a wide-range of possible abnormalities or hazardous items in an x-ray (Wolfe, Alaoui Soce, & Schill, 2017). Although hybrid search is ubiquitous in daily life, relatively little is known about the neural mechanisms that support memory search during search for multiple targets (Ort & Olivers, 2020).

In hybrid search tasks, response time increases linearly *with the log* of the number of search targets stored in memory (memory set size, MSS), suggesting that target verification is highly efficient (Wolfe, 2012). During search for multiple targets, one could imagine that multiple "search templates" are pre-activated in visual working memory (VWM), biasing attention towards target-relevant features in the environment (Desimone & Duncan, 1995). Then, at a post-selection stage, the observer determines whether the attended item matches any of their pre-activated search templates (Ort & Olivers, 2020). This use of VWM would be limited by the capacity of that VWM, which is assumed to be around 4 items (Luck & Vogel, 1997). However, when the number of targets exceeds the small capacity limitations of VWM, target verification was suggested to rely on long-term memory (LTM) recognition (Ort & Olivers, 2020). LTM appears to be essentially free of capacity limitations for targets that are sufficiently learned (Brady, Konkle, Alvarez, & Oliva, 2008). Support for this comes from hybrid search studies, in which observers memorized and search for up to 100 unique targets quite effectively (Wolfe, 2012). Eye-tracking data showed that observers searched the entire display and performed a memory search for each selected item until the actual target was found (Drew, Boettcher, & Wolfe, 2017). Furthermore, performance was not impaired if a concurrent task was used to "fill up" VWM capacity during search (Drew, Boettcher, & Wolfe, 2016). Though search was not impaired, VWM capacity seemed to be reduced by roughly one item when performing a hybrid search task concurrently with a VWM load. This evidence suggests that the target load in hybrid search does not fill up VWM. Rather, VWM may serve as a fixed-capacity conduit that passes a single attended item at a time into LTM for target verification.

While it seems clear that high target loads in hybrid search do fill up VWM, templates in VWM may still guide hybrid search even at large MSSs (Cunningham & Wolfe,

---

[1]University of California San Diego, [2]Radboud Universiteit, [3]The University of New Mexico, [4]Harvard Medical School, Boston, MA, [5]University of Toronto - Mississauga, [6]Sirona Medical Inc., San Francisco, CA
*authors contributed equally

Journal of Cognitive Neuroscience X:Y, pp. 1–23
https://doi.org/10.1162/jocn_a_02256

2014). This would occur, not by having dozens of search templates in VWM, but by having a few "guiding templates" in VWM that would make item selection more efficient while dozens of "target templates" in LTM would be used to determine if a selected item was indeed a target (Wolfe, 2021). For example, imagine a hybrid search task in which your targets are from the same category (e.g., all targets are animals; perhaps, a cat, a cow, and an octopus). A rough representation of the features of animals could guide attention *away* from items that share few visual features with the target category (e.g., flags). These items could be rejected before eliciting memory search. Other items, animals or distractors with animal-like features, would be attentionally selected and would require memory search to determine if they were a member of the memory set (Cunningham & Wolfe, 2014).

ELECTROPHYSIOLOGICAL CORRELATES OF MULTIPLE-TARGET SEARCH

Effects of target load and categorical similarity on top-down attentional processing of items in hybrid search could be marked by the N2 posterior-contralateral (N2pc). The N2pc is a lateralized posterior event-related potential (ERP) occurring around 200 ms after the stimulus' onset, marking spatial attentional selection (Wiegand et al., 2018; Eimer, 1996; Luck & Hillyard, 1994b). The N2pc is reliably elicited by both distinct-item and categorical targets (Nako, Wu, Smith, & Eimer, 2014; Wu et al., 2013) and is smaller when observers search for one of two targets, as compared to when looking for a single target (Grubert & Eimer, 2016). A recent study provided behavioral and neural evidence for top–down attentional effects of categorical target templates during hybrid search (Shang, Yeh, Zhao, Wiegand, & Peelen, 2024). Specifically, distractors from the same category elicited a stronger N1 and N2pc than distractors from a different category. While these results highlight the role of feature-based attentional selection in multiple-target search, it remains unclear how the number of targets, and their categorical level, influences VWM and LTM processes following attentional selection.

Further insights about these memory processes in hybrid search may be derived from ERPs indexing post-selective processing. First, VWM might be marked by the Contralateral Delay Activity (CDA), a posterior, lateralized sustained negative ERP that occurs during VWM maintenance ~300–1000 ms after the presentation of to-be-remembered items (Luria, Balaban, Awh, & Vogel, 2016; Vogel & Machizawa, 2004). In change-detection VWM tasks, the CDA increases with the number and resolution of maintained items, reaching an asymptote at participants' individual VWM capacity limit (Wiegand et al., 2014; Fukuda, Awh, & Vogel, 2010; Ikkai, McCollough, & Vogel, 2010; Vogel & Machizawa, 2004). Hence, the CDA is regarded as a neural signature of active storage in VWM. In visual search, a CDA (also called sustained posterior contralateral lateralization, SPCN) follows the N2pc if the task requires VWM maintenance for post-selective target discrimination (Wiegand, Finke, Müller, & Töllner, 2013; Hilimire, Mounts, Parks, & Corballis, 2011; Brisson & Jolicoeur, 2008). The CDA has also been used as a neural signature of pre-activating attentional templates in VWM using simple stimuli and small set sizes of one to three targets (Grubert, Carlisle, & Eimer, 2016; Gunseli, Meeter, & Olivers, 2014; Carlisle, Arita, Pardo, & Woodman, 2011; Luria & Vogel, 2011). However, how the CDA varies with LTM target load during memory search for multiple targets has not yet been investigated.

Second, ERP correlates of LTM for large sets of learned items have been studied extensively in the recognition memory literature. Here, previously learned "old" probe items are known to elicit two dissociable ERP components, compared to new items. First, the FN400 occurs 300–500 ms after the probe. It is fronto-centrally distributed and less negative (i.e. more positive) for old items than for new items. The FN400 is thought to reflect familiarity-based recognition, a fast-acting process that does not reflect retrieval of qualitative or contextual details about the encoding event (Rugg & Curran, 2007; Curran & Cleary, 2003; Curran, 2000). Second, the Late Positive Complex (LPC) occurs around 500 ms after the probe. It is centro-parietally distributed, and more positive for old than for new items. The LPC is thought to reflect recollection, the more effortful retrieval process for the recognized item along with its context of occurrence (Rugg & Curran, 2007; Curran, 2000). The FN400 and LPC are elicited for different memoranda, including visual objects (Curran & Cleary, 2003). Thus, the FN400 and/or LPC may also serve as markers of LTM retrieval during target verification in hybrid search.

THE PRESENT STUDY

The goal of the present study was to better understand the contributions of different processes during multiple-target search (Ort & Olivers, 2020). In three experiments, we measured ERP correlates of attentional processing (N2pc), VWM (CDA), and LTM recognition (FN400 and LPC) in a modified hybrid search task, where visual set size was constrained to only one task-relevant item per trial, to isolate the effects of searching through memory rather than searching through space. Note that this task is thereby similar to old/new recognition paradigms commonly used in the LTM literature (Curran, 2000). Here, the term *memory search* refers to the retrieval of target representations from LTM to match them against the attended items in the visual display. The processes contributing to memory search, and their hypothesized ERP correlates, are illustrated in Figure 1. Across all experiments, we varied the MSS, that is, the number of targets observers would search for. One of the targets was present in half of the trials and a distractor was displayed in the other half. In Experiments 1 and 2, observers searched

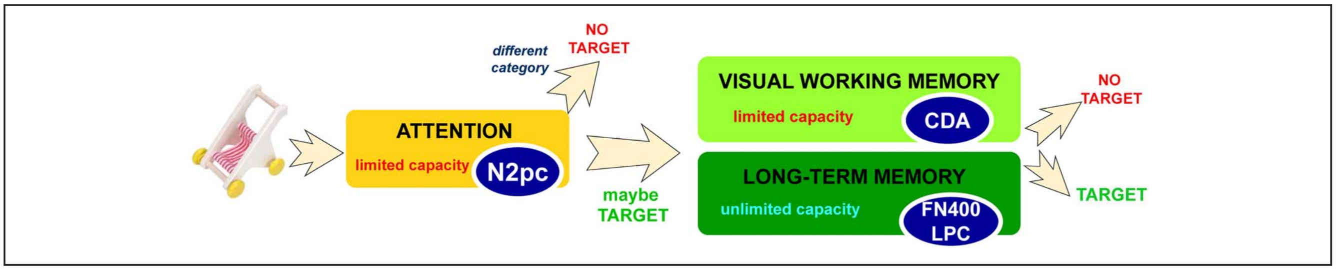

Figure 1. Schematic illustration of processes involved in memory search and their hypothesized ERP correlates. We expected that no memory search would be elicited if search items can be rejected at an attentional stage based on categorical features (Shang et al., 2024); thus, items dissimilar to any of the targets would not be loaded into VWM or LTM. We assumed that the CDA, indexing visual working memory (VWM) load (Vogel & Machizawa, 2004), would increase with memory set size (MSS) up to the VWM capacity-limit of ~4 items. By contrast, we hypothesized that the FN400 and/or LPC, marking the strength of long-term memory (LTM)-based recognition (Rugg & Curran, 2007), would vary with MSS beyond 4 items.

for sets of 1–64 distinct, realistic objects. We tested whether and how the CDA, FN400, and LPC would vary with MSS, marking the contributions of VWM and LTM processes for target loads within and beyond the capacity limitations of VWM. In Experiment 3, observers searched for 2 or 16 target objects either among distractors of the same, a similar, or dissimilar category. We tested whether target dissimilar items would indeed be rejected before eliciting a memory search (Cunningham & Wolfe, 2014), as marked by the N2pc and ERP correlates of memory search established in Experiments 1 and 2.

EXPERIMENT 1

In Experiment 1, observers searched for either 1, 2, 4, or 8 distinct target objects to distinguish between MSS effects on ERPs within and beyond the capacity limitations of VWM. We analyzed response times and accuracy, together with the N2pc, CDA, FN400 and LPC as a function of MSS and target presence. First, we expected that an N2pc should be elicited, marking spatial attentional processing of target items. Note that spatial attention to the target location was already pre-cued by color in the present design; thus, the N2pc cannot be interpreted to reflect template-based guidance to target features in the presence of distractors. Rather, the N2pc may reflect attentional processing of an item *at* the attended location, which is expected to be easier if one, or a few, search templates can be activated compared to when many targets are in the memory set. Second, if a limited number of pre-activated search templates within capacity limitations of VWM supports target verification (Ort & Olivers, 2020), we may expect the CDA to increase with MSS but plateau at MSS4. Alternatively, if VWM indeed only passes the attended item into LTM for target verification (Drew et al., 2016, 2017), a CDA should be present, but should not be sensitive to MSS. For target sets exceeding the small capacity limitations of VWM, we expected that target verification would rely largely on LTM recognition (Ort & Olivers, 2020), which would be reflected in a modulation of the FN400 and LPC by target presence (i.e., old/new difference) and target load beyond 4 items.

Method

Participants

Twenty-six participants from the University of Utah participated in Experiment 1 for course credit or $15.00 an hour. Four participants had artifact rejection-rates that exceeded our exclusion criteria (> 30% of trials) and the full EEG session was inadvertently not recorded for two participants, leaving a total of 20 participants in the final dataset (15 female / 5 male, average age: 22.4 years, age range: 18 to 34 years). A sample size of 20 was pre-registered (https://osf.io/saz82). The study was approved by the University of Utah's Institutional Review Board, and all participants provided informed written consent.

Memory Search Task

Participants completed a memory search task for sets of 1, 2, 4, or 8 (MSS) unique target items (Figure 2), which were sampled randomly for each participant from Brady et al. (2008). For each MSS, participants completed a task block with three phases: memorization, recognition test, and memory search. During the memorization phase, participants memorized the items in their target list as they were displayed one at a time at the center of the screen for 3000 ms. Next, participants completed an old/new recognition memory test to ensure the targets had been memorized. The target objects and an equal number of foils were presented one at a time at the center of the screen, and participants indicated whether or not the object was a member of the target memory set. In order to proceed to the next phase of the experiment, participants were required to pass the recognition test with a score of 100% twice. Participants were allowed five attempts to complete the testing phase, and no participant failed to meet this criterion. Finally, participants completed the critical memory search phase. At the beginning of each search block, participants were instructed to attend to the item cued in either a blue or green frame (counterbalanced across participants). At the beginning of each trial, a blank screen containing only a fixation cross was displayed for 300–400 ms. To measure lateralized ERP components, the attended item was displayed to the left or right side of

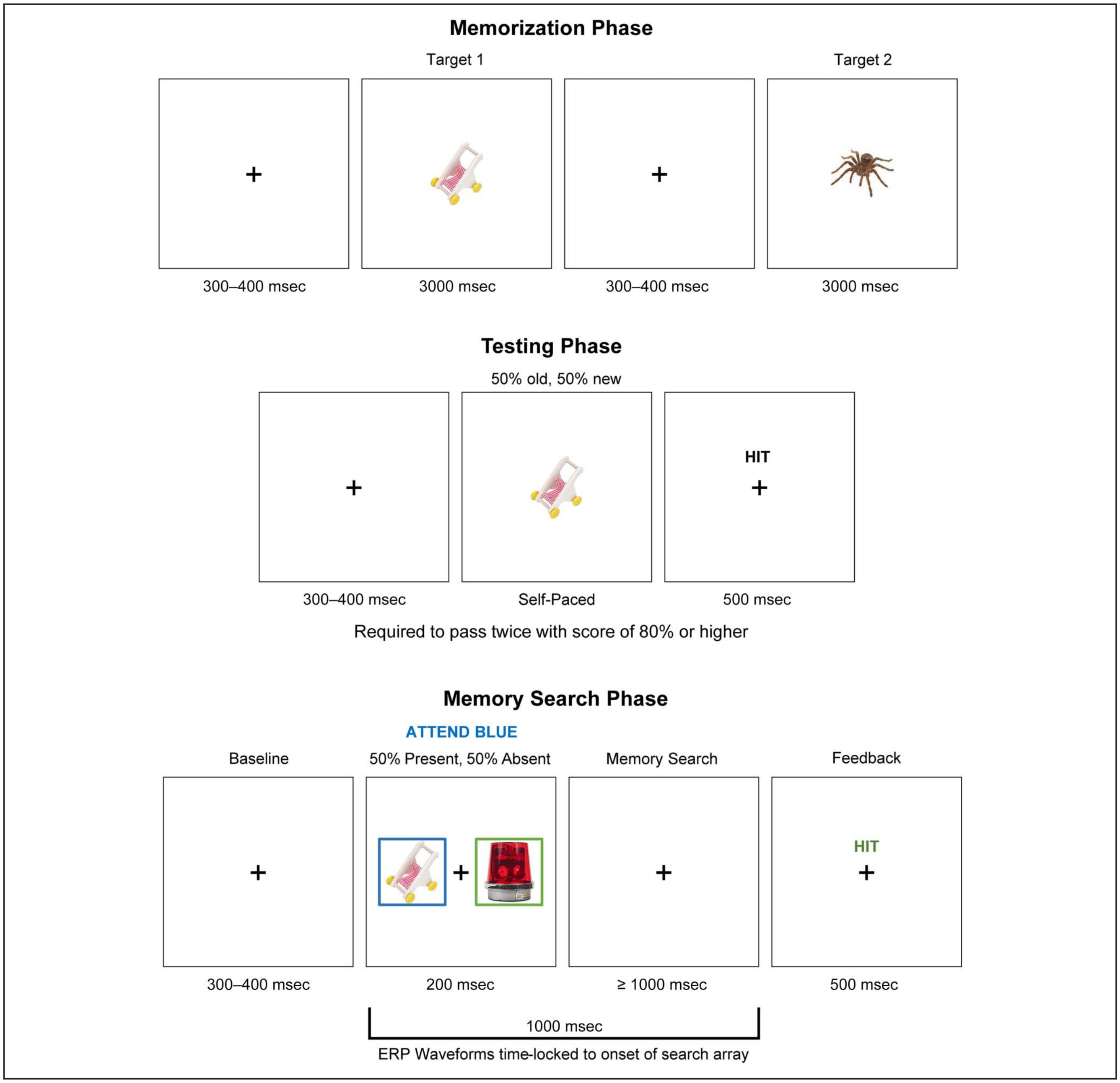

Figure 2. Experimental design for Experiments 1 (Memory Set Size, MSS 1, 2, 4, & 8) and 2 (MSS 2, 4, 16, & 64). The passing rate of the memory test phase was 100% in Experiment 1, and 80% or higher in Experiment 2.

the fixation cross in the task-relevant color (e.g., blue), and an unattended item was shown on the opposite side of the fixation screen in the task-irrelevant color (e.g., green). The objects appeared for 200 ms, followed by a response interval where only the fixation cross was visible. Participants indicated whether or not the object was a member of the target set using the 'F' or 'J' keys. For each MSS, there were 400 trials (50% target present, 50% target absent). Note that the item status (target absent / present) and the hemifield in which the cued stimulus was presented were counterbalanced. Thus, the response key (side of motor response) and visually attended side were not systematically related. Consequently, any response-related lateralization was independent from visual lateralization in the ERP (Wiegand et al., 2013). The response screen was displayed for either 1000 ms or until participants made a response. Finally, participants were shown feedback on their performance until they pressed the space bar to start the next trial. The order of blocks was counterbalanced.

EEG Procedure

EEG data was recorded at 500 Hz with Brain Products' ActiCap and ActiCHamp system using 32 active electrodes from the International 10/20 system. The data were referenced online to the average of the left and right mastoids. Electrode impedance was reduced to 15 kOhms or lower at each electrode site prior to recording Twenty-eight channels besides mastoids were placed according to the 10–20 system with 2 additional placed on the outer canthi bilaterally to recorde HEOG and Fpz serving as the ground. The channels included Fz, Cz, Pz, and Oz, as well as Fp1, F7, F3, FC5, FC1, T7, C3, CP5, CP1, P7, P3, PO7, and contralateral homologues.

EEG Analysis

The EEG data was processed offline using the EEGLAB (Delorme & Makeig, 2004) and ERPLAB (Lopez-Calderon & Luck, 2014) toolboxes for MATLAB. First, a high-pass filter at 0.01 Hz with a half-amplitude cutoff was applied to the raw EEG data. Next, the data was epoched from −200–1000 ms and trials with artifacts were removed from the analysis (see below). The stimulus-locked ERP waveforms were time-locked to the onset of the search display and extended for 1000 ms. The response-locked waveforms (see Appendix) were time-locked to the participant's response and went backward in time for 1000 ms. Waveforms were baseline-corrected to the 200 ms prior to the onset of the search display. For eye-movement detection, two electrodes were placed ~1 cm from the external canthi of each eye, and an HEOG channel was created offline by subtracting the left and right eye channels (Bipolar HEOG = HEOG left – HEOG right). A step-function was then applied to the HEOG channel in order to detect eye-movements (threshold 40 μV, see also Williams & Drew, 2021). Blinks and other large artifacts were detected using a moving-window function (threshold 140 μV) applied to the frontal electrodes (Fp1/2) above each eye. Individual thresholds were adjusted for each participant as needed in order to increase the signal to noise ratio. On average, 9.7% of trials were excluded for the stimulus-locked waveforms and 9.3% of trials were excluded for the response-locked waveforms (see Appendix). We checked the averaged residual activity in the HEOG, which was minimal (< 0.1 degree). In addition, the data was visually inspected for each participant to ensure the appropriateness of our rejection criteria. Finally, we applied a low-pass Butterworth filter with a half-amplitude cutoff of 30 Hz for plotting purposes only. The statistical analyses were performed on the data prior to the application of the low-pass filter.

Event-related Potentials

The selection of electrodes and time windows was pre-registered (https://osf.io/saz82) and in accordance with previous studies of the N2pc and CDA (Drew, Williams, Jones, & Luria, 2018; Vogel & Machizawa, 2004; Eimer, 1996; Luck & Hillyard, 1994), and FN400 and LPC, specifically for using object images (Drew et al., 2018; Küper & Zimmer, 2018; Küper, Groh-Bordin, Zimmer, & Ecker, 2012).

For the lateralized components, the N2pc and CDA, we created a contralateral-ipsilateral difference wave relative to the attended side of the display using the average of electrodes P07/08 and P7/8. The mean amplitude of the N2pc was measured between 200–300 ms. The mean amplitude of the CDA was measured between 300–1000 ms. We also analyzed the response-locked CDA (−300 to 0 ms) to ensure any observed differences in mean amplitude between conditions were not driven by differences in response-time (Ankaoua & Luria, 2023; Williams & Drew, 2021), which are reported in the Appendix. Overall, the results of the analyses on the response-locked CDA mirrored our findings in the stimulus-locked analyses.

For the non-lateralized components, the FN400 and LPC, we used the average of electrodes Fz, Cz, Pz, F3/4, C3/4, and P3/4. The mean amplitude of the FN400 was measured between 300–500 ms relative to stimulus onset, and the mean amplitude of the LPC was measured from 500–800 ms relative to stimulus onset. In addition, we created an old-new difference wave by subtracting the target absent trials from the target present trials and measured the mean amplitude and 50% fractional area latency of the waveform in the time window 300–800 ms (i.e., the point at which the area reached 50% of the total area between 300–800 ms).

Statistical Analysis

The analyses for Experiment 1 were pre-registered (https://osf.io/saz82), except the analyses of the fractional area latency. The ERP and response-time analyses were performed for correct trials only. For each dependent measure, we conducted a 2 (present, absent) by 4 (MSSs 1, 2, 4, 8) repeated-measures ANOVA. This analysis differed from the pre-registered plan to evaluate MSS effects using separate one-way ANOVAs for target present and target absent trials to look for interactions between target presence/absence and MSS. None of the MSS effects substantively differed if we instead use the pre-registered analysis. In addition to frequentist statistics, we computed Bayes Factors for each analysis in order to quantify the degree of evidence for the alternative relative to the null hypothesis (BF_{10}). Bayes Factors greater than 3 were considered sufficient evidence for the alternative hypothesis, and Bayes Factors less than 1/3 were considered sufficient evidence for the null hypothesis (Jeffreys, 1998). Effects and interactions were followed up with Tukey's tests for multiple comparisons.

Results

Response Time and Errors

Response time increased with MSS, $F(3, 57) = 7.66$, $p < .001$, $\eta^2 = .29$, $BF_{10} = 410531.8$ (Figure 3A, Table 1).

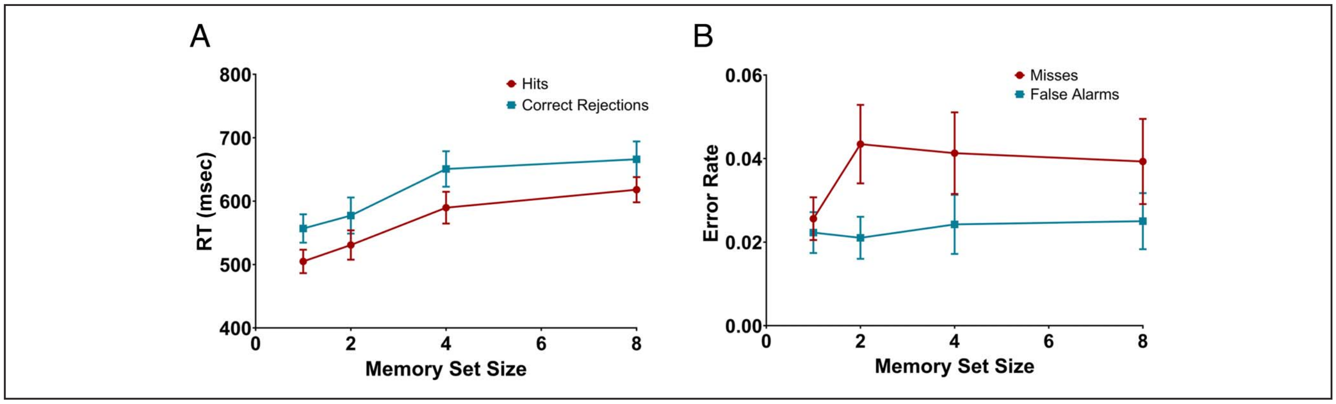

Figure 3. Behavioral data for Experiment 1: (A) response time. (B) error rate (proportion of incorrect responses). Error bars represent the standard error of the mean throughout the manuscript.

Significant differences were found for MSS 1 vs. 4, 1 vs. 8, and 2 vs. 8, all $p < .01$. Responses were faster for target present trials than target absent trials, $F(1, 19) = 32.18$, $p < .001$, $\eta^2 = .63$, $BF_{10} = 30.65$. The MSS by target presence interaction was not statistically significant, $F(3, 57) = 0.38$, p = .77, $\eta^2 = .02$, $BF_{10} = .075$.

Accuracy did not vary with MSS, $F(3, 57) = 0.82$, p = .49, $\eta^2 = .04$, $BF_{10} = .14$ (Figure 3B, Table 1). Participants made more errors in target present trials than target absent trials (i.e., more misses than false alarms), $F(1, 19) = 22.53$, $p < .001$, $\eta^2 = .54$, $BF_{10} = 137.25$. The MSS by target presence interaction was not statistically significant, $F(3, 57) = 2.57$, p = .06, $\eta^{2\,=}$ .12, $BF_{10} = .25$.

N2pc and CDA

Lateralized ERPs are shown in Figure 4. The N2pc, presumably marking attentional processing of the selected item, was larger for target present than target absent trials, $F(1, 19) = 113.6$, $p < .001$, $\eta^2 = .86$, $BF_{10} = 2.33e+27$. While the visual inspection of the grand-averaged N2pc suggests that the amplitude was higher for MSSs 1 and 2 compared to 4 and 8 in target present trials, neither the main effect of MSS, nor the MSS by target presence interaction were statistically significant, both $F(3, 57) < 1.27$, $p > .10$, $\eta^2 < .07$, $BF_{10} < .06$ (Figure 4C, Table 1).

We assumed that post-selective processing would be marked by the CDA. Specifically, we expected that the CDA amplitude would increase with MSS up to the capacity limitation of VWM of ~4 items, but not further. Indeed, the amplitude of the CDA varied with MSS, $F(3, 57) = 4.15$, p = .01, $\eta^2 = .18$, $BF_{10} = .57$. Furthermore, target present trials had a larger CDA amplitude than target absent trials, $F(1, 19) = 9.8$, p = .006, $\eta^2 = .34$, $BF_{10} = 3518.45$. The MSS by target presence interaction was also significant, $F(3, 57) = 6.16$, p = .001, $\eta^2 = .24$, $BF_{10} = 9.15$, reflecting that the CDA did not vary with MSS for target absent trials, p = .53. Interestingly, the target present CDA amplitude was larger at MSS 8 than MSS 4, p = .009. Thus, different from our prediction, the CDA increased beyond the usual capacity limitations of VWM, rather than plateauing after MSS 4 (Figure 4D, Table 1).

Table 1. Descriptive Statistics of Experiment 1

MSS		RT	ER	N2pc	CDA	FN400	LPC	ONE-A	ONE-L
1	TP	504.9 (80.5)	2.5 (2.2)	−2.0 (1.3)	−0.7 (0.7)	0.1 (2.1)	0.5 (1.9)	2.5 (1.6)	500.2 (50.4)
	TA	556.8 (97.6)	2.2 (2.1)	−0.5 (0.7)	−0.6 (0.7)	−3.5 (3.3)	−1.3 (2.3)		
2	TP	530.8 (101.4)	4.3 (4.3)	−2.4 (1.4)	−0.9 (0.9)	−0.3 (2.9)	1.4 (1.9)	2.8 (1.2)	495.9 (47.9)
	TA	577.3 (124.1)	2.1 (2.2)	−0.3 (0.8)	−0.5 (0.6)	−4.2 (2.9)	−0.7 (1.7)		
4	TP	589.7 (109.3)	4.1 (4.2)	−2.1 (1.0)	−1.2 (1.0)	−1.7 (3.3)	1.6 (1.9)	3.3 (1.2)	520.3 (42.4)
	TA	650.7 (122.2)	2.4 (3.1)	−0.3 (0.7)	−0.6 (0.7)	−5.5 (2.7)	−1.3 (1.9)		
8	TP	618.0 (86.4)	3.9 (4.4)	−1.9 (1.0)	−1.6 (1.1)	−2.4 (3.4)	1.4 (1.8)	3.0 (1.5)	540.9 (53.0)
	TA	666.0 (122.5)	2.5 (2.9)	−0.2 (0.7)	−0.5 (0.9)	−5.6 (3.1)	−1.5 (2.1)		

Means and standard deviations (in parentheses) are reported for behavioral and event-related potential measures in Experiments 1, for the different memory set sizes (MSS) and target present (TP) and target absent (TA) trials. Reaction time (RT) and the latency of the old/new effect (ONE-L) is reported in milliseconds. Errors (ER) are reported in percent errors. Amplitudes of the N2 posterior-contralateral (N2pc), contralateral delay activity (CDA), FN400, late posterior complex (LPC), and old/new effect (ONE-L) are reported in μV, ONEs are computed by subtracting TA from TP trials.

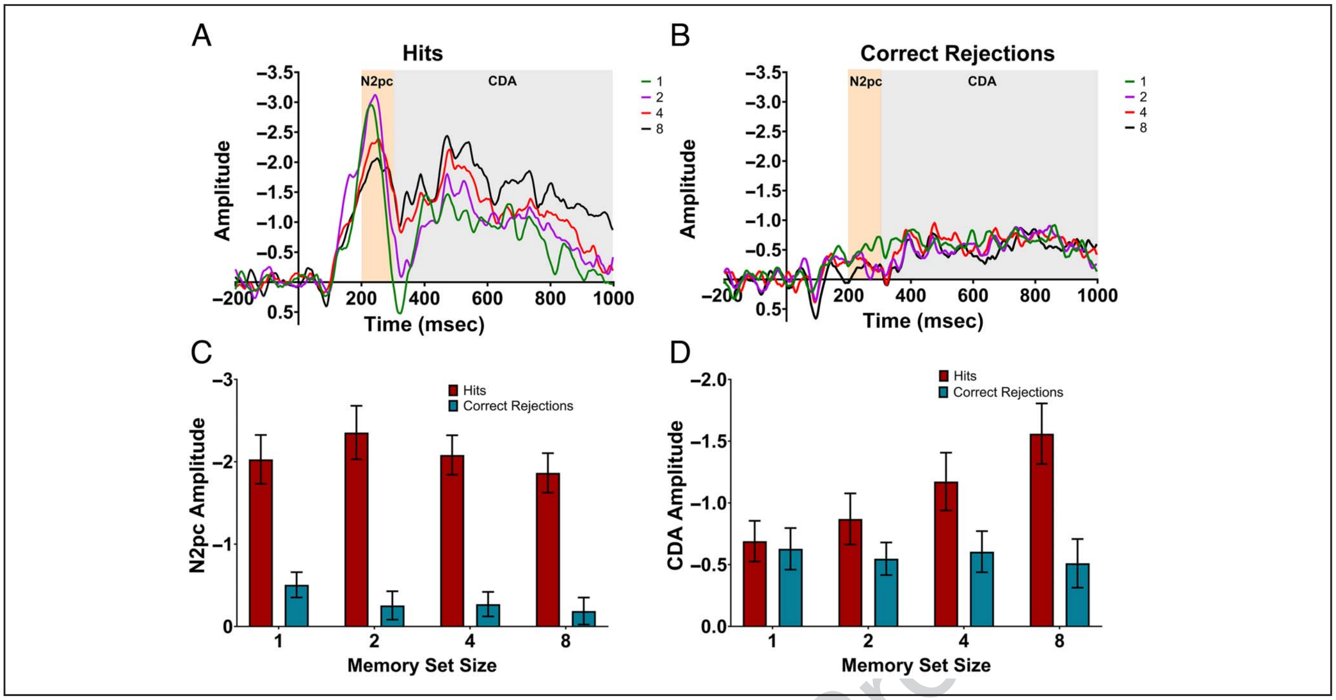

Figure 4. Lateralized ERP data for Experiment 1: (A) contralateral-ipsilateral waveforms for target present trials. (B) contralateral-ipsilateral waveforms for target absent trials. (C) mean N2pc amplitude. (D) mean CDA amplitude. ERPs are averaged across electrodes P7/P8 and PO7/PO8.

FN400, LPC, and Old-new Effect

The non-lateralized ERPs are shown in Figure 5. If memory search relies on LTM recognition, we would expect that the FN400 and LPC would be more positive if observers recognize "old" targets compared to when being presented with a "new" distractor item. Consistent with this prediction, both the FN400 and the LPC were more positive in target present than target absent trials, $F(1, 19) = 95.9, p < .001, \eta^2 = .83, \mathrm{BF}_{10} = 4.51\mathrm{e}{+}20$ and $F(1, 19) = 104.3, p < .001, \eta^2 = .85, \mathrm{BF}_{10} = 1.55\mathrm{e}{+}18$.

Furthermore, we expected the recognition signal to be stronger when a smaller set of targets needed to be

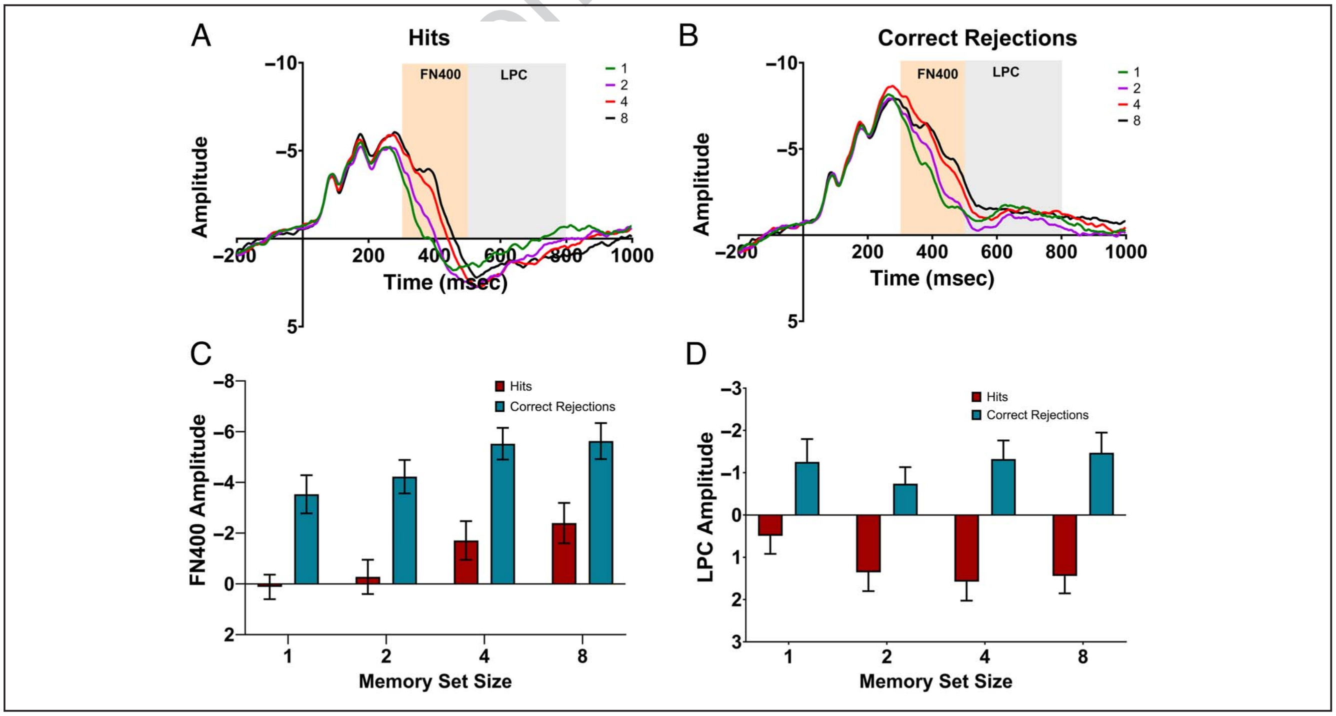

Figure 5. Non-lateralized ERP data for Experiment 1: (A) non-lateralized waveforms for target present trials. (B) non-lateralized waveforms for target absent trials. (C) mean FN400 amplitude. (D) mean LPC amplitude. ERPs are averaged across electrodes F3, Fz, F4, C3, Cz, C4, P3, Pz, P4.

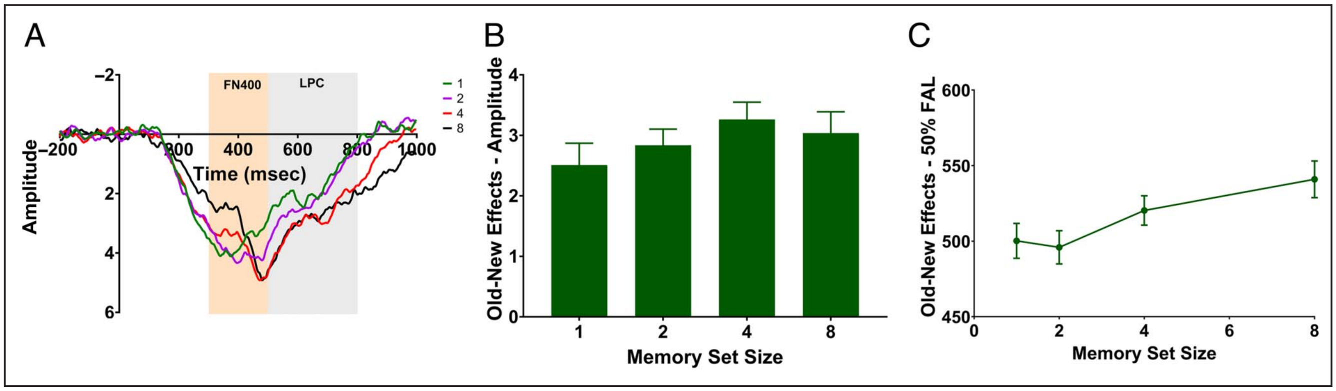

Figure 6. Old/new difference waves for Experiment 1: (A) old/new difference waveforms (B) mean amplitude of the old/new difference wave (C) mean fractional area latency of the old/new difference wave. ERPs are averaged across electrodes F3, Fz, F4, C3, Cz, C4, P3, Pz, P4.

distinguished from distractors compared to a larger set. This was true for the FN400. The FN400 varied with MSS, $F(3, 57) = 16.22, p < .001, \eta^2 = .46$, $BF_{10} = 192.23$ (Figure 5C, Table 1) and was significantly smaller (less negative) in amplitude for MSS 1 than 4, $p < .001$, and 8, $p < .001$, and smaller for MSS 2 than 4, $p = .004$, and set size 8, $p < .001$. There were no significant differences between any of the other MSSs, all $p > .05$. The MSS by target presence interaction was not significant, $F(3, 57) = 0.78$, $p = .51$, $\eta^2 = .04$, $BF_{10} = .10$; thus, the FN400 was modulated by MSS across target present and absent trials (i.e., hits and correct rejections). Different from the FN400, the LPC did not vary with MSS, $F(3, 57) = 1.27$, $p = .29$, $\eta^2 = .06$, $BF_{10} = .10$ (Figure 5D, Table 1) across all trials. However, there was a significant MSS by target presence interaction, $F(3, 57) = 5.34$, $p = .003$, $\eta^2 = .22$, $BF_{10} = .54$. Only in target present trials, the LPC was less positive for set size 1 than for set size 2, $p = .007$, and for set size 4, $p = .002$. No significant differences in the LPC between MSSs were found in target absent trials (all $p > .05$).

Old-new difference waves (ERP in response to target-present trials minus ERP in response to target-absent trials across the time windows of the FN400 and LPC) are plotted in Figure 6. The amplitude of the old-new difference wave did not vary with MSS, $F(3, 57) = 1.60$, $p = .20$, $\eta^2 = .08$, $BF_{10} = .37$ (Figure 6B, Table 1). However, the fractional area latency differed between MSSs, $F(3, 57) = 4.93$, $p = .004$, $\eta^2 = .21$, $BF_{10} = 11.77$. MSS 8 had a significantly longer latency than MSS 1, $p = .02$, and MSS 2, $p = .006$ (Figure 6C, Table 1).

Discussion

In Experiment 1, we found that ERP correlates of both VWM and LTM recognition, the CDA and FN400, respectively, varied with target load beyond set sizes of four targets during memory search. To test whether target load already influences attentional processing, we further analyzed the effect of MSS on the N2pc. We found that the N2pc was only apparent for target present trials, suggesting attentional processing of targets but little attentional processing of non-targets. This result is in agreement with recent hybrid search experiments (Lavelle, Luria, & Drew, 2023), showing that feature-based attention can persist for target numbers exceeding VWM capacity. Furthermore, one might have expected that the N2pc would decrease with MSS, given that attentional guidance is assumed to be more effective when the number of search templates is smaller (Ort & Olivers, 2020). Others have shown that the N2pc is smaller when observers search for two targets relative to only one target (Grubert & Eimer, 2016). The

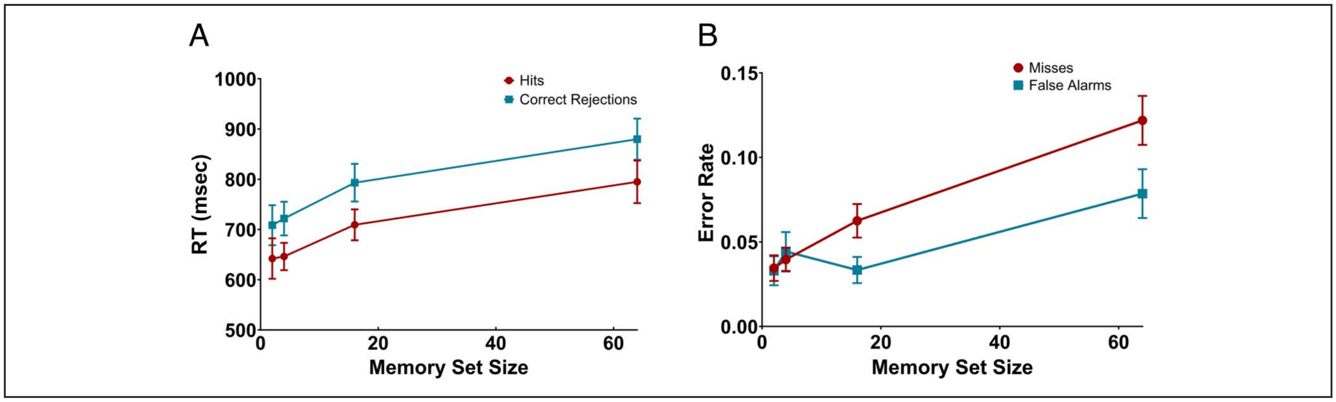

Figure 7. Behavioral data for Experiment 2: (A) response time. (B) error rate (proportion of incorrect responses).

N2pc in the present experiment decreased only numerically, but not statistically significantly, with MSS. Note, however, that the design of the present task minimized the role of guidance by pre-cueing the target location from the outset. Thus, this N2pc does not reflect guidance, but only attentional processing of an item at an already selected location, which may also attenuate the MSS effect.

Interestingly, deviating from our predictions, the CDA significantly increased up to MSS 8 in target present trials. Previous research on hybrid search suggested that VWM usage does not increase with MSS, but rather that VWM passes single, attended items into LTM for comparison to the list of targets (Drew et al., 2016). Furthermore, in VWM studies, the CDA plateaus at VWM load of 3–4 items (Luria et al., 2016; Vogel & Machizawa, 2004). Our results may be interpreted to reflect that not all items of the MSS need to be loaded into VWM during memory search. Presumably, only "suspicious" targets from the memory set that are somewhat similar to the search item may fill up VWM (see Figure 1). The higher the MSS, the more targets will resemble a search item and be loaded into VWM. However, dissimilar targets will not be loaded into VWM; thus, the physical MSS would not be the number of targets actually loaded into VWM. This may also explain why the CDA was smaller and the MSS effect was absent in target absent trials. By using sets of distinct objects, the similarity between targets and distractors is generally small. Thus, most non-targets may have been rejected already at an attentional stage, not eliciting a memory search (Shang et al., 2024).

As we expected, the FN400 and LPC were more positive if a target was present, showing the old/new effect reflecting item recognition (Curran, 2000). Further, there was an effect of MSS on the amplitudes of the FN400 and the latency of the old/new effect. The FN400 is thought to reflect item familiarity and has been related to perceptual matching processes when comparing visually processed information to memory representations (Küper et al., 2012; Zimmer & Ecker, 2010). The effect of target load suggests that the familiarity signal was weaker when more candidate items needed to be retrieved from LTM in order to be compared to the attended item in the search display during memory search. The prolonged latency of the old/new difference waves with larger MSS may reflect the increase in time required for matching more items of the memory set against the attended item. By contrast, the LPC, marking recollection of items and their study context (Norman, Tepe, Nyhus, & Curran, 2008), varied little between MSSs. This suggests that differences between MSS conditions do not result from recollection of the study episode. This is in line with the rather small response time costs of increasing the target load, supporting the assumption that memory search is a fast-acting matching process, which does not rely on time-consuming recollection of each item in the memory set (Nosofsky, Cox, Cao, & Shiffrin, 2014; however, see Wiegand & Wolfe, 2020; Wolfe, Boettcher, Josephs, Cunningham, & Drew, 2015).

EXPERIMENT 2

The findings of Experiment 1 suggest that the involvement of both VWM and LTM increases with target load up to eight target items in memory search. In Experiment 2, we sought to replicate and extend our findings with MSSs up to 64. Observers searched for 2, 4, 16, or 64 targets. We analyzed response times and accuracy, together with the N2pc, CDA, FN400, LPC and old/new effects, as a function of MSS and target presence. First, we expected that an N2pc should be elicited in target present trials as in Experiment 1, marking spatial attentional processing of the target item. Second, we tested whether, as in Experiment 1, the CDA amplitude would increase as MSS exceeded 4, here increasing to 16 and 64 targets. This would suggest an increase in VWM usage with target load, however, not following the item-specific capacity limitations in VWM. Finally, we expected to replicate that the amplitude of the FN400 amplitude and latency of the old/new effect increase with MSS, marking prolonged recognition of targets with growing target sets.

Method

Participants

Thirty-nine observers from the University of Utah participated in the study in exchange for course credit or $15.00 an hour. Like in Experiment 1, we aimed for a sample size of 20 after exclusion, but overshot this goal. Ten participants had artifact rejection-rates that exceeded our exclusion criteria (> 30% of trials) and one participant did not complete the study, leaving 28 participants in the final dataset (10 female/1 non-binary, average age: 22.8 years, age range: 18–36 years). The study was approved by the University of Utah's Institutional Review Board, and all participants provided informed consent.

Memory Search Task

In Experiment 2, the MSSs were 2, 4, 16, and 64. Participants were required to pass the testing phase twice with a score of 80%. There were 296 trials for each MSS (50% target present, 50% target absent) in the memory search phase, and the task-relevant color was counterbalanced between experiment blocks rather than participants. Otherwise, the experimental design was identical to Experiment 1.

EEG Procedure and Analysis

The EEG procedure and analyses were identical to Experiment 1. On average, 14.9% of trials were rejected in the stimulus-locked waveforms and 12.6% of trials were rejected in the response-locked waveforms (see Appendix).

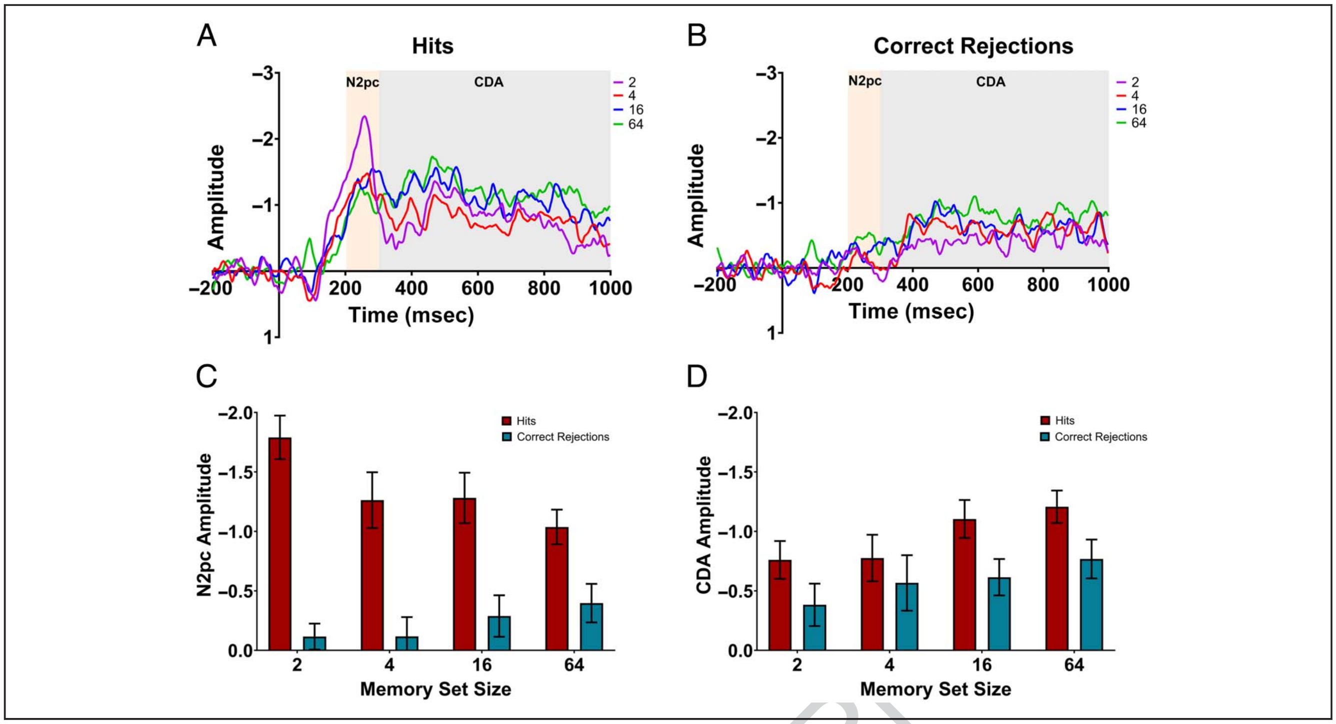

Figure 8. Lateralized ERP data for Experiment 2: (A) contralateral-ipsilateral waveforms for target present trials. (B) contralateral-ipsilateral waveforms for target absent trials. (C) mean N2pc amplitude. (D) mean CDA amplitude. ERPs are averaged across electrodes P7/P8 and PO7/PO8.

Results

Response Time and Errors

Response time increased with MSS, $F(3, 81) = 15.54, p < .001$, $\eta_p^2 = .37$, $BF_{10} = 7.01e{+}10$ (Figure 8A, Table 2). Response time differed between all MSSs, $p < .0001$, except between 2 vs. 4, $p > .71$. Response time was significantly faster in target present trials than target absent trials, $F(1, 27) = 33.85$, $p < .001$, $\eta_p^2 = .55$, $BF_{10} = 2582.51$. The MSS by target presence interaction was not significant, $F(3, 81) = 0.49$, $p = .69$, $\eta_p^2 = .02$, $BF_{10} = .06$.

Accuracy decreased with MSS, $F(3, 81) = 30.66$, $p < .001$, $\eta_p^2 = .53$, $BF_{10} = 3.06e{+}16$ (Figure 8B, Table 2). Observers made significantly more errors when MSS 64 was compared to the other MSSs, $p < .0001$. Accuracy was significantly lower for target present trials than target absent trials, $F(1, 27) = 7.5$, $p = .01$, $\eta_p^2 = .22$, $BF_{10} = 2.59$; observers showed more misses than false alarms. The

Table 2. Descriptive Statistics of Experiment 2

MSS		RT	ER	N2pc	CDA	FN400	LPC	ONE-A	ONE-L
2	TP	642.2 (209.4)	3.4 (3.9)	−1.8 (1.0)	−0.8 (0.8)	−0.8 (3.4)	2.5 (2.2)	3.7 (1.7)	531.6 (63.0)
	TA	708.5 (206.5)	2.4 (2.7)	−0.1 (0.8)	−0.4 (0.9)	−5.0 (3.7)	−0.8 (2.1)		
4	TP	646.2 (141.9	3.9 (3.7)	−1.3 (1.2)	−0.8 (1.0)	−2.9 (3.9)	2.0 (2.4)	3.5 (1.8)	559.8 (73.0)
	TA	721.7 (174.5)	3.4 (3.9)	−0.1 (0.8)	−0.6 (1.2)	−6.4 (4.3)	−1.4 (2.7)		
16	TP	709.3 (159.6)	6.3 (5.1)	−1.3 (1.1)	−1.1 (0.8)	−3.0 (3.7)	2.4 (2.5)	3.2 (1.3)	601.1 (82.9)
	TA	793.2 (194.8)	3.8 (3.7)	−0.3 (0.9)	−0.6 (0.8)	−6.1 (4.0)	−0.9 (2.7)		
64	TP	794.8 (220.5)	12.2 (7.5)	−1.0 (0.8)	−1.2 (0.7)	−4.5 (4.0)	0.8 (3.4)	2.7 (1.3)	794.8 (220.5)
	TA	879.7 (213.4)	9.1 (7.8)	−0.4 (0.8)	−0.8 (0.8)	−6.5 (4.2)	−2.4 (3.2)		

Means and standard deviations (in parentheses) are reported for behavioral and event-related potential measures in Experiments 2, for the different memory set sizes (MSS) and target present (TP) and target absent (TA) trials. Reaction time (RT) and the latency of the old/new effect (ONE-L) are reported in milliseconds. Errors (ER) are reported in percent errors. Amplitudes of the N2 posterior-contralateral (N2pc), contralateral delay activity (CDA), FN400, late posterior complex (LPC), and old/new effect (ONE-A) are reported in μV, ONEs are computed by subtracting TA from TP trials.

MSS by target presence interaction was not significant, $F(3, 81) = 2.04$, p = .11, $\eta_p^2 = .07$, $BF_{10} = .17$.

N2pc and CDA

Lateralized ERPs are shown in Figure 8. As in Experiment 1, the N2pc was prominent in target present, but not in target absent trials, $F(1, 27) = 104.6$, $p < .001$, $\eta_p^2 = .79$, $BF_{10} = 3.31e{+}17$. Again, the main effect of MSS on the N2pc amplitude was not significant, $F(3, 81) = 1.11$, p = .35, $\eta_p^2 = .04$, $BF_{10} = .06$ (Figure 8C, Table 2). However, the MSS by target presence interaction was statistically significant, $F(3, 81) = 5.23$, p = .002, $\eta_p^2 = .16$, $BF_{10} = 6.55$. For target present trials, the N2pc was larger (i.e., more negative) for MSS 2 than MSSs 4, 16, and 64, all $p < .05$. No differences between MSSs were significant for target absent trials, all $p > .05$.

In line with our findings in Experiment 1, the grand-averaged ERPs suggest that the CDA continuously increased with MSSs. Indeed, the CDA was significantly modulated by MSS, $F(3, 81) = 4.18$, p = .008, $\eta_p^2 = .13$, $BF_{10} = .86$ (Figure 8D, Table 2), but post hoc tests showed that the CDA was only significantly smaller (i.e., less negative) for MSS 2 than for MSS 64, p = .01. There were no significant differences between any of the other MSSs, all $p > .05$, after correcting for multiple comparisons. The CDA amplitude was larger for target present than target absent trials, $F(1, 27) = 10.77$, p = .003, $\eta_p^2 = .29$, $BF_{10} = 42.54$. Different from Experiment 1, the MSS by target presence interaction was not statistically significant, $F(3, 81) = 0.27$, p = .85, $\eta_p^2 = .01$, $BF_{10} = .07$; thus, the effect of target load was comparable for target present and target absent trials.

FN400, LPC, and Old/New Effects

Non-lateralized ERPs are shown in Figure 9. As expected, we replicated the old/new difference, that is, more positive amplitudes for "old" targets compared to "new" distractor items, in the FN400, $(1, 27) = 209.4$, $p < .001$, $\eta_p^2 = .89$, $BF_{10} = 9.43e{+}22$, and the LPC, $F(1, 27) = 138.5$, $p < .001$, $\eta_p^2 = .83$, $BF_{10} = 4.92e{+}27$. Furthermore, we replicated the FN400 modulation by MSS, $F(3, 81) = 15.31$, $p < .001$, $BF_{10} = 30888.11$ (Figure 9C, Table 2). The FN400 was significantly smaller for MSS 2 than MSSs 4, 16, and 64, all $p < .001$. There were no significant differences between any of the other MSSs, all $p > .05$, after correcting for multiple comparisons. The set size by target presence interaction was statistically significant, $F(3, 81) = 15.72$, $p < .001$, $\eta_p^2 = .37$, $BF_{10} = 10.97$. In target present trials, the FN400 increased with MSS; all comparisons were significant, $p < .0001$, except the difference between 4 and 16. For target absent trials, the FN400 amplitude at MSS 2 was significantly smaller than for 4, 16, and 64, $p < .0001$, but the other MSSs did not differ, $p > .62$.

The LPC amplitude also varied with MSS, $F(3, 81) = 6.99$, $p < .001$, $\eta_p^2 = .21$, $BF_{10} = 12.49$ (Figure 9D, Table 2). Specifically, the LPC was significantly smaller (i.e., less positive) for the largest MSS of 64 compared to 2, 4, and 16, all $p < .05$. There were no significant

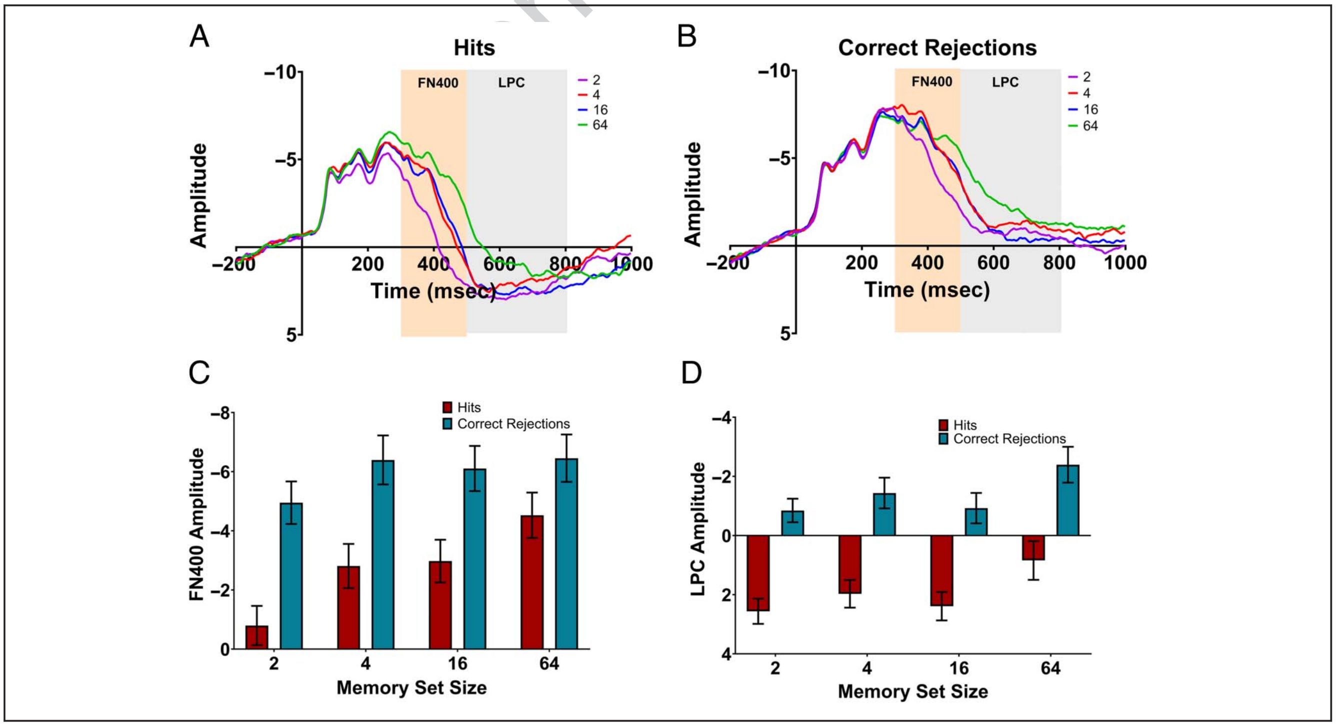

Figure 9. Non-lateralized ERP data for Experiment 2: (A) non-lateralized waveforms for target present trials. (B) non-lateralized waveforms for target absent trials. (C) mean FN400 amplitude. (D) mean LPC amplitude. ERPs are averaged across electrodes F3, Fz, F4, C3, Cz, C4, P3, Pz, P4.

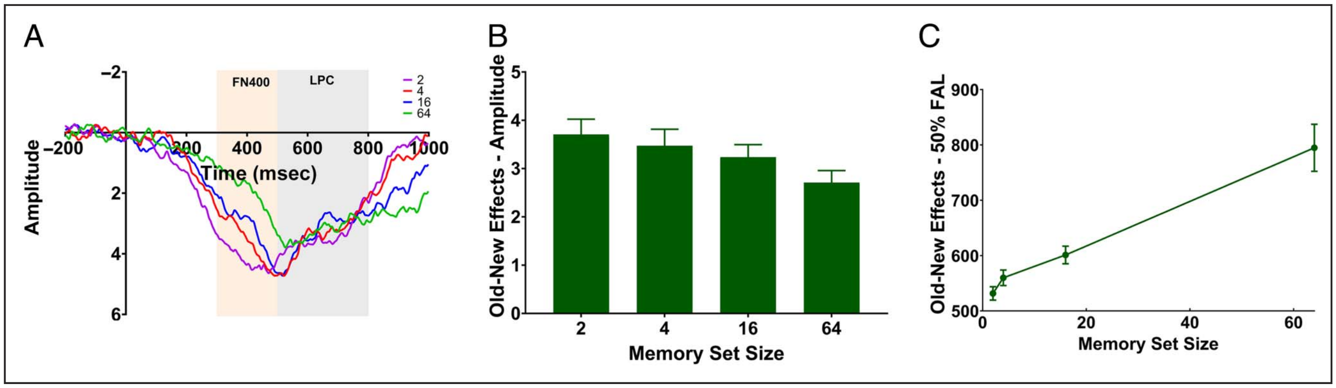

Figure 10. Old/new difference waves for Experiment 2: (A) old/new difference waveforms (B) mean amplitude of the old/new difference wave (C) mean fractional area latency of the old/new difference wave. ERPs are averaged across electrodes F3, Fz, F4, C3, Cz, C4, P3, Pz, P4.

differences between any of the other MSSs, all $p > .05$. The MSS by target presence interaction was not significant, $F(3, 81) = 0.13$, $p = .94$, $\eta_p^2 = .004$, $BF_{10} = .06$.

Old-new difference waves are shown in Figure 10. The mean amplitude varied with MSS, $F(3, 81) = 4.37$, $p = .007$, $\eta_p^2 = .14$, $BF_{10} = 5.70$ (Figure 10B, Table 2). MSS 64 had a significantly smaller amplitude than MSS 2, $p = .03$. None of the other MSS comparisons were statistically significant, all $p > .05$. Replicating the finding of Experiment 1, MSS modulated the fractional area latency of the old-new difference waves, $F(3, 81) = 12.11, p < .001$, $\eta_p^2 = .31$, $BF_{10} = 16566.40$. Fractional area latency was significantly shorter for MSS 2 than 16, $p = .03$, and 64, $p < .001$, shorter for MSS 4 than 64, $p = .02$, and shorter for MSS 16 than 64, $p < .001$ (Figure 10C, Table 2).

Discussion

In Experiment 2, we largely replicated findings from Experiment 1 using larger MSSs. As before, the N2pc was only observable in target present trials, even for larger MSSs, supporting that some selective attention to relevant features even exists with (very) large numbers of targets (Lavelle et al., 2023). Furthermore, the N2pc was larger for MSS 2 compared to all other MSSs. This might reflect that, at set size 2, observers could pre-activate two, relatively distinct, search templates, which facilitates attentional processing of the target if the MSSs is within the limits of VWM capacity (Ort & Olivers, 2020). At higher set sizes, perhaps only one or a few templates were preactivated.

Interestingly, we replicated that the CDA increased with MSS; here with target load up to MSS 64, well beyond the item-capacity limits of VWM. This supports our assumption that the involvement of VWM during memory search cannot be understood simply in terms of a limited number of "spots" being filled with any target from LTM. Different from Experiment 1, the CDA also increased with MSS in target absent trials. As mentioned above, due to the perceptual and semantic overlap between targets and non-targets at larger MSS, more "suspicious" non-targets will elicit a memory search.

Finally, also in accordance with Experiment 1, the FN400 and the latency of the old/new effect increased gradually with MSS. In addition, we found the LPC and old/new effect to be reduced for the largest set size of 64. These findings suggest that the recognition signal decreases, and occurs later, with increasing target load during memory search. Presumably, familiarity-based recognition is weaker because larger perceptual and semantic overlap between targets and non-targets in large target sets of distinct, highly variable objects cause interference. This may also lead to a higher decision threshold for making old/new judgments, as more perceptual evidence needs to be accumulated to classify an item as target or non-target. Furthermore, higher error rates in the conditions with 64 targets suggest that participants may also have guessed correctly on a number of trials without a reliable recognition signal being present.

EXPERIMENT 3

Having established ERP correlates of memory search in Experiments 1 and 2, we sought to use these measures to test predictions from Cunningham and Wolfe's (2014) model of hybrid search in Experiment 3. A key assumption of that model is that memory search is not needed if items can be ruled out by attention to their categorical features (Shang et al., 2024). For example, when searching for multiple cats simultaneously, attention will be guided towards cats and items with similar features (e.g., dogs). These 'lure' items will elicit a memory search and will only be rejected during the target verification stage. In contrast, items that share few visual features with the cat target items (e.g., tables) will be rejected at an earlier stage and will not require a memory search (see Figure 1).

The CDA results in Experiment 1 and 2 suggest that VWM may indeed be filled selectively, only with "suspicious" targets from the memory set that are somewhat similar to the search item. In Experiment 3, we manipulated the MSS and the similarity between the search item and target category in order to test this assumption directly. Observers searched for 2 or 16 target objects from one category (e.g. cats), either among distractors of the same (other cats), a similar (dogs), or dissimilar (tables) category. If target-dissimilar items can be rejected before eliciting a memory search (Cunningham & Wolfe, 2014), we would expect that MSS influences the ERP correlates of memory search only if the attended item is of the same or a similar category as the target items. Specifically, we expected that targets and target-similar non-targets would elicit a larger N2pc than target-dissimilar non-targets, marking attentional processing of items with categorical target features (Shang et al., 2024). Furthermore, we expected that the CDA and FN400 would increase with MSS for targets and non-targets with similar categorical features, marking memory search elicited by "suspicious" items, while the CDA and FN400 would not vary with MSS if observers attend to non-targets from a dissimilar category.

Method

Participants

Twenty-five participants from the University of Utah participated in the study for course credit or $15.00 an hour. Three participants exceeded the artifact rejection-rate threshold (> 30% of trials) and two participants experienced equipment failure, leaving a total of 20 participants in the final dataset (13 female, 7 male, average age: 20.1 years, age range: 18–34 years), matching the pre-registered sample size (https://osf.io/tmurs).

Memory Search Task

The task structure of Experiment 3 was the same as Experiments 1 and 2 (Figure 11). However, in Experiment 3, participants memorized target sets containing objects from the same category. The MSSs were 2 and 16, and the possible target categories were dogs, cats, tables, and dressers. Each participant had one target set from the animal categories (e.g., dogs) and one target set from the furniture categories (e.g., tables) in separate blocks. Target categories and MSSs were counterbalanced across participants, such that there were an equal number of each category per MSS. In the old/new recognition memory test, the foil items were from the same category as the target set for that block. During the memory search phase, the attended item varied between four conditions: a target (e.g., a target dog), a non-target from the target category (e.g., a non-target dog), a non-target from a similar category (e.g., a non-target cat), and a non-target from a dissimilar category (e.g., a non-target dresser). Each participant completed 174 trials per condition.

EEG Procedure and Analysis

The EEG procedure and analyses were identical to Experiments 1 and 2. 12.9% of trials were excluded from the stimulus-locked analysis and 9.6% of trials were excluded from the response-locked analysis (see Appendix).

Statistical Analysis

The analyses were pre-registered: (https://osf.io/tmurs). After collecting data from the first few participants, it became apparent that a categorical hybrid search task at MSS 64 would result in insufficient correct trials for our EEG analyses. Thus, we modified our pre-registration to include smaller MSSs and more trials. We performed a 2 (MSS 2, 16) × 4 (Conditions 1–4) repeated measures ANOVA on each of the dependent variables. Main effects and interactions were followed-up with Tukey's multiple comparisons.

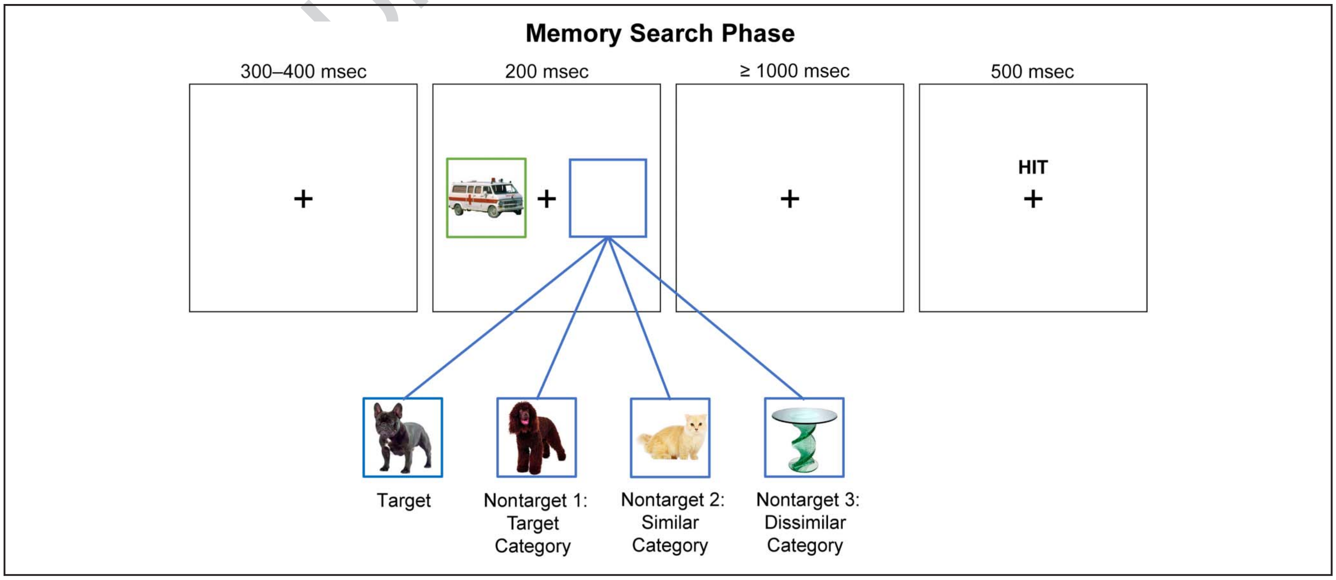

Figure 11. Experimental design for Experiment 3.

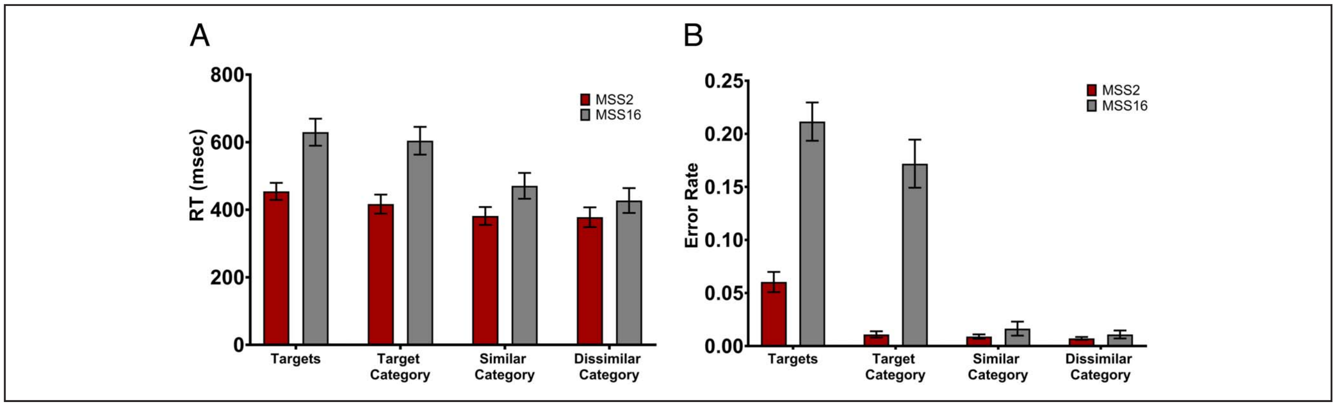

Figure 12. Behavioral data for Experiment 3: (A) response time. (B) error rate (proportion of incorrect responses).

Results

Response Time and Errors

Response time decreased as the search item became more dissimilar to the target, $F(3, 57) = 52.00, p < .001, \eta_p^2 = .73$, $BF_{10} = 234501.56$ (Figure 12A, Table 3). Specifically, response times were longer for targets than non-targets from the similar category, $p < .001$, and non-targets from the dissimilar category, $p < .001$. Response times were slower for non-targets from the target category than non-targets from the similar category, $p < .001$, and non-targets from the dissimilar category, $p < .001$. Response times were overall faster for MSS 2 than 16, $F(1, 19) = 16.87$, p = .001, $\eta_p^2 = .47$, $BF_{10} = 3.24e+8$, and the search item condition by MSS interaction was significant, $F(3, 57) = 31.69, p < .001$, $\eta_p^2 = .63$, $BF_{10} = 40.02$. The MSS effect was significant for the target condition, $p < .001$, and the target category condition, $p < .001$, and marginally significant for the similar category, p = .06. The MSS effect was not significant for the dissimilar category, p = .43.

Accuracy increased as the search item became more dissimilar to the target, $F(3, 57) = 75.90, p < .001, \eta_p^2 = .80$, $BF_{10} = 2.14e+12$ (Figure 12B, Table 3). Specifically, participants made more errors in classifying targets compared to non-targets from the target category, $p < .001$, non-targets from the similar category, $p < .001$, and non-targets from the dissimilar category, $p < .001$. Error rates were also higher for non-targets from the target category than for non-targets from the similar category, $p < .001$, and for non-targets from the dissimilar category, $p < .001$. Accuracy did not significantly differ between non-targets from the similar and the dissimilar category, p = .99. Accuracy was overall significantly better for MSS 2 than MSS 16, $F(1, 19) = 65.12, p < .001, \eta_p^2 = .77$, $BF_{10} = 4.96e+6$, and the search item condition by MSS interaction was significant, $F(3, 57) = 53.69, p < .001, \eta_p^2 = .73$, $BF_{10} = 6.03e+14$, such that the effects of categorical similarity were more pronounced at the higher MSS. MSS effects were significant for the target condition, $p < .001$, and the condition with distractors from the target category, $p < .001$.

Table 3. Descriptive Statistics of Experiment 3

MSS		RT	ER	N2pc	CDA	FN400	LPC	ONE-A	ONE-L
2	Target	454.4 (110.2)	6.0 (4.1)	−2.4 (1.9)	−0.9 (0.8)	2.6 (4.0)	4.7 (3.3)		
	Same	416.7 (123.0)	1.0 (1.3)	−1.0 (1.3)	−0.8 (0.9)	−2.3 (3.4)	0.4 (3.1)	4.5 (1.9)	526.5 (40.5)
	Similar	381.5 (115.7)	0.9 (0.9)	−1.0 (1.1)	−0.4 (0.6)	−2.2 (3.2)	0.4 (2.8)	4.5 (2.3)	532.0 (42.7)
	Dissimilar	377.8 (127.3)	0.7 (0.5)	−0.9 (1.0)	−0.5 (0.7)	−3.3 (2.9)	0.0 (2.8)	5.2 (2.2)	523.9 (44.0)
16	Target	629.6 (174.5)	21.1 (7.8)	−3.2 (2.0)	−2.0 (1.0)	−0.3 (4.5)	3.7 (4.3)		
	Same	604.2 (179.2)	17.2 (9.9)	−2.3 (1.9)	−1.6 (1.0)	−2.8 (3.4)	−0.2 (3.9)	3.3 (1.5)	565.6 (35.9)
	Similar	470.8 (166.3)	1.6 (2.8)	−1.9 (1.4)	−0.8 (0.9)	−3.2 (3.6)	0.6 (3.2)	3.0 (2.5)	543.4 (60.0)
	Dissimilar	427.4 (159.8)	1.0 (1.6)	−0.6 (1.1)	−0.2 (0.5)	−4.1 (3.5)	−0.5 (3.3)	4.0 (3.4)	561.3 (61.2)

Means and standard deviations (in parentheses) are reported for all behavioral and event-related potential measures in Experiments 3, for the two memory set sizes (MSS) and in response to targets, non-targets from the target category (same), non-targets from a similar category, and non-targets from a dissimilar category. Reaction time (RT) and the latency of the old/new effect (ONE-L) are reported in milliseconds. Errors (ER) are reported in percent errors. Amplitudes of the N2 posterior-contralateral (N2pc), contralateral delay activity (CDA), FN400, late posterior complex (LPC), and old/new effect (ONE-A) are reported in μV, Note that ONEs are computed by subtracting trials in which non-targets (same, similar, dissimilar) are presented from trials in which a target is presented.

MSS effects were not significant for the condition with distractors from a target-similar category, p = .7, or the target-dissimilar category, p = .82.

N2pc and CDA

The lateralized ERPs are plotted in Figure 13. We expected that the N2pc would mark attention to categorical target features (Shang et al., 2024). Indeed, the amplitude of the N2pc significantly decreased (became less negative) as the search item became more dissimilar to the target, $F(3, 57) = 16.20$, $p < .001$, $\eta_p^2 = .46$, $BF_{10} = 3.37e{+}8$ (Figure 13C, Table 3). The N2pc was larger for targets relative to all non-targets, all $p < .003$. However, the N2pc was also larger for non-targets from the target category than for non-targets from the dissimilar category, p = .02. Furthermore, the N2pc was significantly larger for MSS 16 than MSS 2, $F(1, 19) = 17.16$, p = .001, $\eta_p^2 = .47$, $BF_{10} = 13.02$, and the condition by MSS interaction was significant, $F(3, 57) = 7.28$, p = .002, $\eta_p^2 = .28$, $BF_{10} = 4.16$. The MSS effect was significant for the targets, p = .02, non-targets from the target category, $p < .001$, and non-targets from the similar condition, p = .002. MSS effects were not significant for non-targets from the dissimilar category, p = .72.

As in Experiments 1 and 2, the CDA amplitude (Figure 13D, Table 3) increased significantly with MSS, $F(1, 19) = 19.92, p < .001$, $\eta_p^2 = .51$, $BF_{10} = 69.90$, confirming again that the involvement of VWM increases with target load during memory search. If VWM involvement during this memory search is indeed restricted to items that share categorical features with the target (Cunningham & Wolfe, 2014), we would expect that the CDA to be more pronounced and to vary with MSS only for targets and non-targets that resemble the target. Indeed, the CDA varied with similarity between the attended item and the target category, $F(3, 57) = 21.52$, $p < .001$, $\eta_p^2 = .53$, $BF_{10} = 8.93e5$, and increased gradually with similarity to the target: CDA amplitudes were larger for targets than non-targets from the similar category, $p < .001$, and the dissimilar category, $p < .001$, and larger for non-targets from the target category than non-targets from the similar category, p = .004, and non-targets from the dissimilar category, $p < .001$. The CDA amplitude did not significantly differ between non-targets from the similar and dissimilar category, p = .30. Furthermore, the condition by MSS interaction was significant, $F(3, 57) = 7.23$, p = .004, $\eta_p^2 = .28$, $BF_{10} = 413.39$. The difference between MSS 2 and 16 was significant for targets, $p < .0001$, and items from the target category, p = .0005, but not for items from a similar, p = .14, and dissimilar category, p = .18.

FN400, LPC and Old/New Effects

Non-lateralized ERPs are shown in Figure 14. As in Experiment 1 and 2, the amplitude of the FN400 and LPC differed between "old" target items and "new" non-target items, $F(3, 57) = 38.71$, $p < .001$, $\eta_p^2 = .67$, $BF_{10} = 9.92e{+}14$ and $F(3, 57) = 55.18$, $p < .001$, $\eta_p^2 = .74$, $BF_{10} = 2.29e{+}16$. The FN400 and LPC were more positive for targets

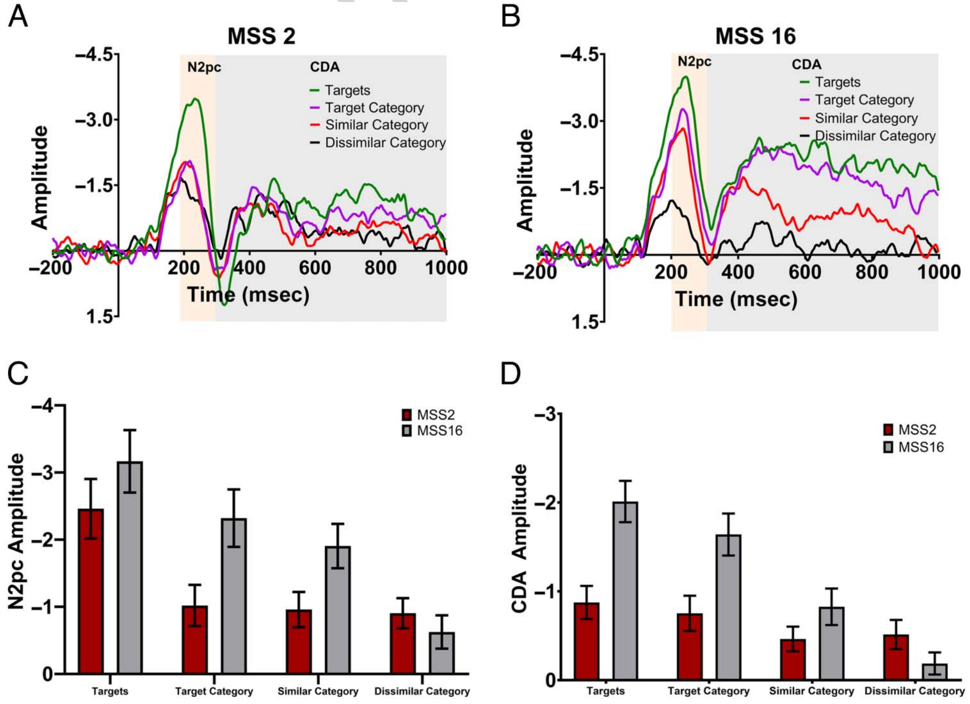

Figure 13. Lateralized ERP data for Experiment 3: (A) contralateral-ipsilateral waveforms for memory set size (MSS) 2. (B) contralateral-ipsilateral waveforms for MSS 16. (C) mean N2pc amplitude. (D) mean CDA amplitude. ERPs are averaged across electrodes P7/P8 and PO7/PO8.

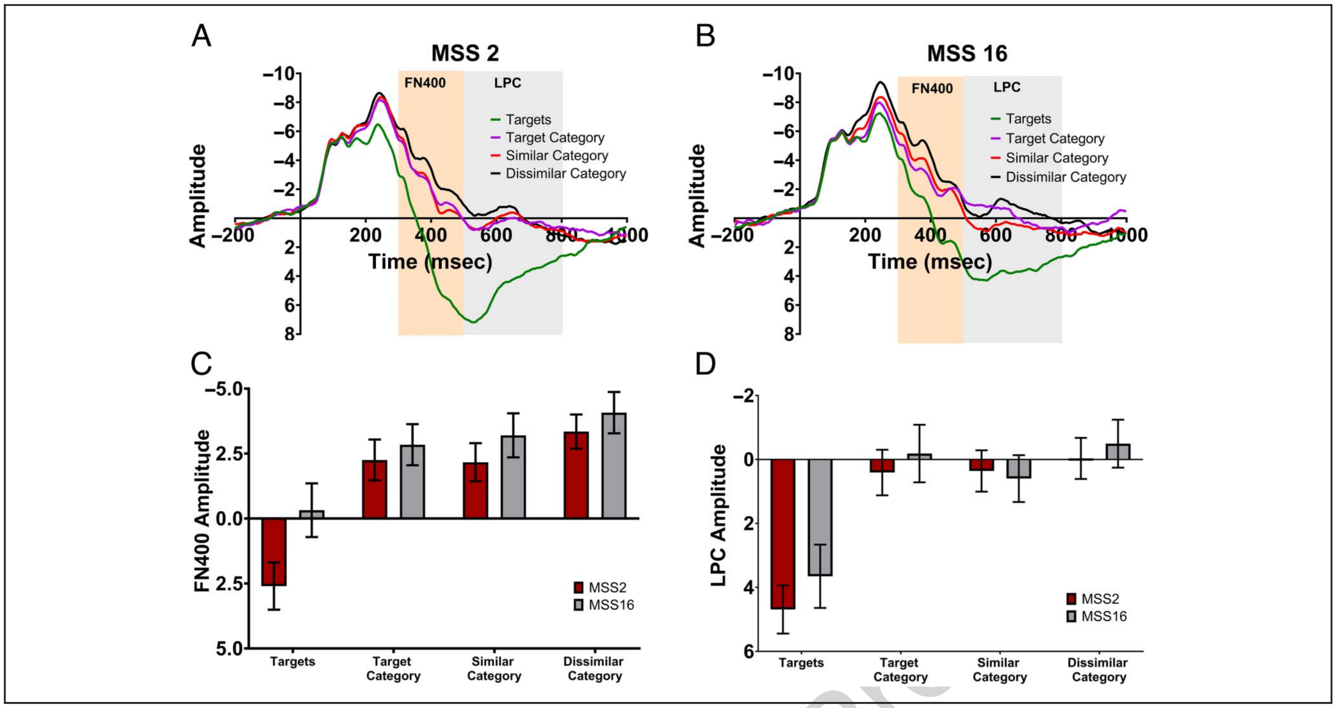

Figure 14. Non-lateralized ERP data for Experiment 3: (A) non-lateralized waveforms for memory set size (MSS) 2. (B) non-lateralized waveforms for MSS 16. (C) mean FN400 amplitude. (D) mean LPC amplitude. ERPs are averaged across electrodes F3, Fz, F4, C3, Cz, C4, P3, Pz, P4.

compared to all non-target conditions $p < .0001$; while the FN400 and LPC between non-targets of the target category, similar category, and dissimilar category did not differ significantly from each other after adjusting for multiple comparisons, all $p > .08$.

Assuming that memory search is only elicited if the attended item shared categorical features with the target category; we expected the FN400 to vary with MSS if the search item is a member of the target or target-similar category, but not if the search item is a member of a target-dissimilar category. Indeed, the FN400 varied with MSS, $F(1, 19) = 8.35$, p = .009, $\eta_p^2 = .31$, $\mathrm{BF}_{10} = 8.32$ (Figure 14C, Table 3) and MSS interacted significantly with condition, $F(3, 57) = 5.76$, p = .002, $\eta_p^2 = .23$, $\mathrm{BF}_{10} = 1.40$. However, the FN400 was only significantly larger (i.e., more negative) for MSS 16 than MSS 2 in the target condition, p = .001, while the MSS effect was not significant for any of the non-target conditions, all $p > .05$. In accordance with the findings from Experiment 1, the LPC amplitude did not significantly vary with MSS, $F(1, 19) = 0.91$, p = .35, $\eta_p^2 = .05$, $\mathrm{BF}_{10} = .30$. and neither was the condition by MSS interaction significant, $F(3, 57) = 0.93$, p = .41, $\eta_p^2 = .05$, $\mathrm{BF}_{10} = .13$ (Figure 14D, Table 3).

The old-new difference waves are shown in Figure 15. Its amplitude decreased from MSS 2 to 16, $F(1, 19) = 4.79$, p = .01, $\eta_p^2 = .20$, $\mathrm{BF}_{10} = 136.16$, and differed significantly between non-target conditions, $F(2, 38) = 6.07$, p = .02, $\eta_p^2 = .24$, $\mathrm{BF}_{10} = .47$ (Figure 15C, Table 3). Specifically, the old-new difference was larger for target-dissimilar non-targets than for target-similar non-targets, p = .02, and marginally significantly larger for target-dissimilar non-targets than for non-targets of the target category, p = .06. The condition by MSS interaction was not significant, $F(2, 38) = 0.14$, p = .87, $\eta_p^2 = .01$, $\mathrm{BF}_{10} = .14$. Also the fractional area latency of the old-new effects significantly differed between non-target conditions, $F(1, 19) = 9.07$, p = .007, $\eta_p^2 = .32$, $\mathrm{BF}_{10} = .12$ (Figure 15D), but none of the post-hoc tests survived the correction for multiple comparisons, all $p > 10$. The main effect of MSS on the fractional area latency was not significant, $F(2, 38) = 0.80$, p = .46, $\eta_p^2 = .04$, $\mathrm{BF}_{10} = 1402.10$; however, MSS interacted with the non-target condition, $F(2, 38) = 3.44$, p = .04, $\eta_p^2 = .15$, $\mathrm{BF}_{10} = .51$. Specifically, the old-new effect occurred later for MSS 16 than 2 for non-targets of the target category, $p < .001$, and for non-targets from a dissimilar category, $p < .001$.

Discussion

In Experiment 3, we replicated the finding that the FN400 and the CDA are sensitive to target load during memory search. In addition, our results support the assumption that target-similar items are attentionally processed and elicit a memory search, while target-dissimilar items can be rejected prior to engaging in memory search (Shang et al., 2024; Lavelle et al., 2023; Cunningham & Wolfe, 2014): Observers responded more slowly and less accurately to search items that are more similar to the target; particularly when the MSS was larger. These behavioral costs of target similarity and MSS were mirrored in

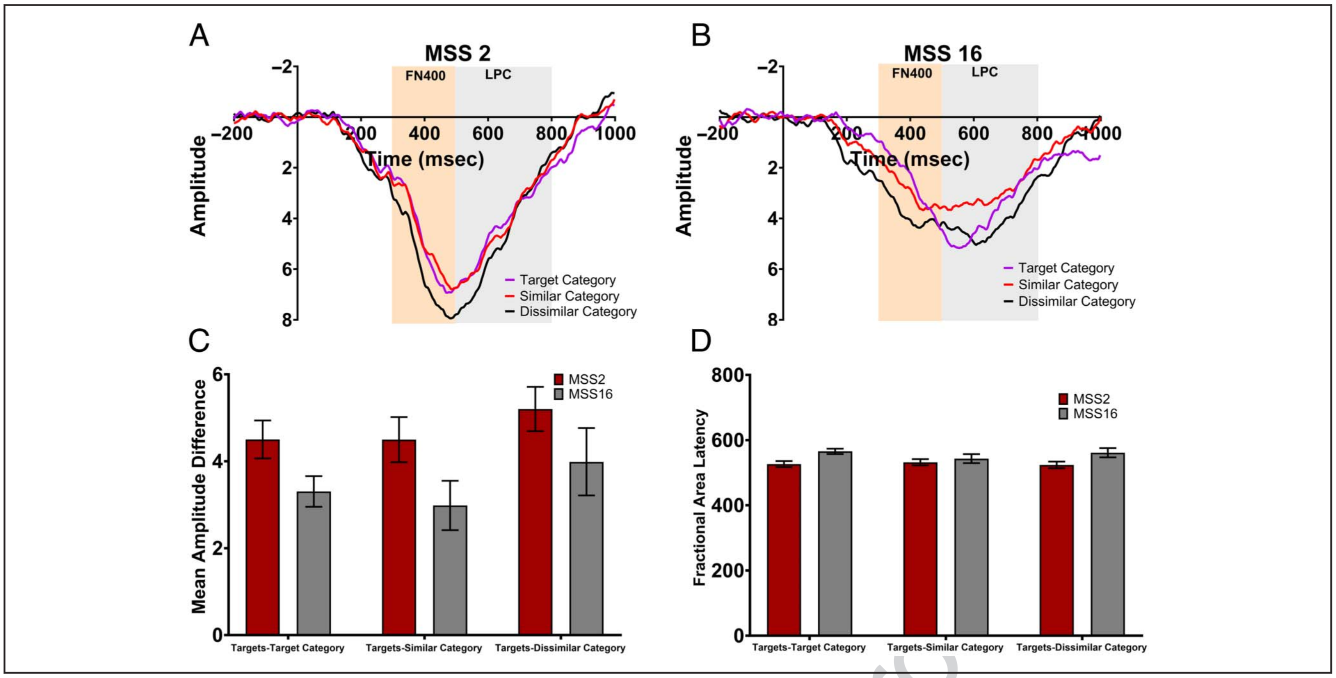

Figure 15. Old/new difference waves for Experiment 3: (A) old/new difference waveforms for memory set size (MSS) 2 (B) old/new difference waveforms for MSS 16 (C) mean amplitude of the old/new difference wave (D) mean fractional area latency of the old/new difference wave. ERPs are averaged across electrodes F3, Fz, F4, C3, Cz, C4, P3, Pz, P4.

multiple ERP variations: First, the N2pc was larger in response to search items of the target category than in response to search items from a target-dissimilar category. The N2pc also increased with MSS for targets and non-targets of the target category, but not if the search item was from a target-dissimilar category. This finding is in line with, and extends, recent results (Shang et al., 2024) showing that attentional processing is enhanced for items that match features of the target category. This effect depends on the number of targets in memory: At smaller MSSs, observers could effectively prevent items from similar categories from being attentionally processed, just as they prevented dissimilar items from being processed. At larger MSSs, target-similar non-target items were more likely to be attentionally processed. Previous studies have shown that search for targets defined on a superordinate-category level (e.g., animals, clothing) is slower (Cunningham & Wolfe, 2014), and that the N2pc is reduced and delayed compared to search for distinct targets (Nako et al., 2014; Wu et al., 2013). In the present study, however, where targets were all from the same the basic-category level, categorical status provides a strong means for feature-based attentional selection and processing.

Second, the CDA increased with MSS only if the search item was a target or a non-target from the same target category, which indicates that items with target-matching categorical features activate VWM resources during memory search, while categorically different items put less (or no) load on VWM. At the same time, the FN400 increased with MSS if the search item was a target and the old/new effect gradually decreased as the target-similarity of the non-target and MSS increased. Presumably, the LTM-based recognition is more difficult under high target-distractor similarity and target load due to perceptual and conceptual interference (Konkle, Brady, Alvarez, & Oliva, 2010). Distractors of the target category may elicit a certain degree of "lure" familiarity, specifically if the MSS is larger.

GENERAL DISCUSSION

In the present study, we identified multiple EEG correlates of processes underlying visual memory search by manipulating target load (MSS) and categorical similarity of targets and distractors. In Experiments 1 and 2, observers searched for sets of 1–64 distinct object images. We assumed that VWM load, as indexed by the CDA (Luria et al., 2016), would increase with memory load of small target sets within the capacity limitations of VWM of ~4 items. With MSSs significantly larger than 4, we expected a ceiling effect in the CDA. We expected that search through larger memory sets would rely on LTM recognition processes (Ort & Olivers, 2020), as indexed by the FN400, LPC, and old/new effect (Rugg & Curran, 2007; Curran, 2000). In accord with this second set of expectations, we found that the FN400 and old/new effect varied with MSSs up to 64 during memory search, presumably marking increased demands of LTM retrieval. However, our expectations about the CDA and VWM were not met. The CDA also increased with MSS up to 64, far beyond any reasonable estimate of the capacity of VWM. This suggests that VWM resources during memory search cannot be understood as a simple matter of filling up a small number of "slots" with any item of the large target set.

In Experiment 3, we tested whether categorical similarity of targets and non-targets would influence attention, VWM and LTM processes during memory search. We found electrophysiological support for the idea that memory search for targets derived from a single category can be limited to target-similar items (Shang et al., 2024; Cunningham & Wolfe, 2014). First, the N2pc increased with similarity to the target; in particular, when the MSS was large, suggesting that items dissimilar to the target were rejected on an early attentional processing stage, prior to memory search (see also Shang et al., 2024). Next, target load effects on the CDA and FN400 depended on the categorical status of the search item. Indeed, only items that resembled the target elicited a memory search as indexed by the MSS related increase in the CDA and FN400.

The Role of Attention and Visual Working Memory in Memory Search

Our findings provide compelling evidence that attentional and VWM processes contribute to memory search, even if the target memory set is (very) large. First, the N2pc marks how attentional processing of the search items influences memory search. In line with previous studies, the N2pc amplitude decreased with increasing the number of targets (Grubert & Eimer, 2016), significantly in Experiment 2 with 2–64 distinct targets. Presumably, if the observer looks for one out of two targets, it is possible to activate two distinct search templates, with high fidelity, facilitating attentional processing of targets relative to non-targets. When looking for one target out of 8, 16, or 64, however, this search template will be crude so that attentional processing of targets and non-targets is more equal (i.e., less biased to specific features represented in a search template). Experiment 3 further demonstrated that the N2pc was sensitive to the degree of feature overlap between the search item and the target category (Shang et al., 2024). Targets and non-targets from the same category both elicited a strong N2pc. This suggests that attentional processing of target-similar non-targets contributes to the diminishment of recognition accuracy (Konkle et al., 2010) and hybrid search performance (Lavelle et al., 2023), and influences target verification on post-selection processes, as we see in the modulation of the CDA and FN400 discussed below. In fact, the degree of interference during memory search due to "list length", that is, decline of recognition memory performance with increasing number of items in a memory list (Shiffrin & Steyvers, 1997) could also be understood as a (dis-)ability to attentionally prioritize the memorized items, which is easier, or can be done with higher fidelity, if the number of items is smaller and feature overlap is lower.

With regard to the role of VWM in memory search, we originally hypothesized that the CDA, marking VWM capacity limits (Luria et al., 2016), may plateau around MSSs of ~3–4 items (Luck & Vogel, 1997). Memory search was previously proposed to be VWM-based for set sizes within VWM capacity limits, but LTM-based for set sizes beyond VWM capacity limits (Ort & Olivers, 2020). In accordance with this, our previous behavioral data suggested that VWM usage is not MSS-dependent in hybrid search (Drew et al., 2016). However, all experiments showed that the CDA continued to increase with MSS beyond the VWM capacity limit.

It is unlikely that the CDA increase reflects that many more than four target representations are actively hold in VWM in the present task (Ikkai et al., 2010; Vogel & Machizawa, 2004). Rather, the CDA modulation may reflect sustained effort while the search item is matched against multiple target templates retrieved from LTM. Presumably, only those targets that share features with the perceptually presented item will be loaded into VWM, but not *all* targets from the MSS. This explains that the MSS effect on the CDA does not strictly follow the physical set size of the target set and that the CDA is sensitive to the similarity of the non-targets. More generally, the CDA was shown to rise with difficulty of VWM operations in other tasks, such as mental rotation (Ankaoua & Luria, 2023). In the present task, one possibility is that a higher fidelity version of the attended item might be passed into VWM for comparison to the target memory set when target verification becomes more difficult, taking up more resources in VWM (Bays, Wu, & Husain, 2011). The high fidelity comparison process may also explain why distractors are better recognized when they resemble targets and MSS is larger (Lavelle et al., 2023).

Of note, the CDA (and FN400, see below) and its modulations by MSS were observed in target present and, to a smaller degree, in target absent trials. This is in line with the assumption that both targets and non-targets elicit a memory search in which target verification occurs on post-selective processing stages (Ort & Olivers, 2020). In the experiments with distinct objects, the N2pc was pronounced only in target present trials and not modulated by MSS in target absent trials. Presumably, when targets and non-targets are distinct object images, a number of (randomly) target-dissimilar non-targets might have been rejected easily, thus, were little attended and not loaded into VWM (Hilimire et al., 2011), thus did not elicit an N2pc nor CDA. This would lead to lower amplitudes on average in target-absent, as compared to target-present, trials. Given that the N2pc time window is small, the component cannot be reliably detected if the signal is relatively weak. Alternatively, participants may occasionally engage in memory search even when no target is detected on the screen to reassure target absence. Finally, LTM representations might already be activated and then easily rejected when being presented with a target-dissimilar non-target, requiring little attentional processing of the visually presented features, leading to a small CDA in the absence of a distinct N2pc deflection (Vogel & Machizawa, 2004). In Experiment 3, where target–non-target similarity was systematically varied, both, the N2pc and CDA amplitudes were indeed strongly reduced in response to target-dissimilar non-targets items, supporting that target-dissimilar items are not attended nor loaded into VWM.

Importantly, the experiments reported here were designed to examine the memory search component of hybrid search, rather than the attentional selection stage. Observers were only shown one item on the cued side of the lateralized display, greatly reducing the selective attention requirements in the first place. Thus, it remains unclear how the findings in the N2pc in particular might interact with stronger demands on visual selection, such as increasing the visual set size.

The Role of Long-term Memory in Memory Search

The present experiments show that target load during memory search performance modulated the FN400 and old/new effect. This finding provides electrophysiological support for the assumption that target verification in hybrid search tasks with large sets of well-learned target items relies on LTM recognition (Ort & Olivers, 2020; Drew et al., 2017). Accordingly, the decrease in search performance with increasing MSS could be considered a list length effect, which refers to the phenomenon that longer lists of learned, to-be-recognized items are typically associated with poorer recognition memory performance compared to shorter lists (Shiffrin & Steyvers, 1997). Specifically, the FN400, assumed to mark familiarity-based LTM recognition (Rugg & Curran, 2007), was larger for "old" targets than "new" non-targets and increased gradually with MSS. Furthermore, the old/new difference was prolonged and decreased for large MSSs. Familiarity-based recognition is considered to be a fast and relatively automatic process (Curran, 2000), which may explain why search through many items in memory is remarkably quick (Wolfe, 2012). Consistent with this proposal are recent results showing that familiar items have an advantage over novel stimuli (Madrid et al., 2019). Increasing familiarity of non-targets and reducing familiarity of targets can cause some response time costs in hybrid search (Wiegand & Wolfe, 2020; Wolfe et al., 2015). The FN400 and old/new effect modulations suggest that the familiarity signal to verify an item as "old" becomes weaker and slower the more targets are in the memory set. Presumably, longer lists of distinct target objects cause more interference during the memory search due to larger conceptual and perceptual overlap between the target set and non-targets (Konkle et al., 2010). This is also supported by the findings of Experiment 3, in which the old/new difference between targets and target-similar non-targets is smaller than between targets and target-dissimilar non-targets. Categorical similarity is also associated with conceptual and perceptual overlap (Küper & Zimmer, 2018; Küper et al., 2012) that weakens the discriminative strength of the familiarity signal between targets and target-similar non-targets during memory search.

In contrast to familiarity-based recognition, recollection-based recognition is more time consuming, effortful, and less susceptible to interference by item similarity (Yonelinas, 2002). Behavioral studies have shown that hybrid search remains efficient even if targets cannot be identified based on a familiarity signal alone and a form of rapid recollection may support recognition (Guild, Cripps, Anderson, & Al-Aidroos, 2014). In the present study, we found that the LPC was little modulated by MSS, suggesting that recollection-based retrieval, while likely contributing to target recognition in hybrid search (Wiegand & Wolfe, 2020; Wolfe et al., 2015), remains largely undisturbed by the growing target load.

Conclusions

In recent years, our understanding of search behavior has moved toward increasingly complex and realistic tasks (e.g., radiology; Turoman, Tivadar, Retsa, Murray, & Matusz, 2021; Wolfe, 2021; Nartker, Alaoui-Soce, & Wolfe, 2020). In our daily routines, we often perform "hybrid" visual and memory search; that is, the simultaneous search through the environment for multiple-items stored in memory (Wolfe, 2012). One real-world example of hybrid search is searching through your mental shopping list while looking for the items on your list at the grocery store (Wiegand, Wolfe, Maes, & Kessels, 2024; Boettcher, Drew, & Wolfe, 2018). Although the role of VWM and LTM processes has been studied in relatively simple visual search tasks with one or a few search targets (e.g., Woodman et al., 2007), it was unclear how they act in more complex hybrid search tasks, with possibly hundreds of targets, held in memory. By manipulating target load and categorical similarity of targets and distractors, we identified ERP correlates of multiple processes that contribute to complex search tasks. Our findings highlight that attentional processing of target features, marked by the N2pc, precedes and can control memory search. Following this, VWM and LTM, marked by the CDA and FN400, are engaged concomitantly, even if the number of targets is high.

APPENDIX: ANALYSES OF THE RESPONSE-LOCKED CDA

We analyzed the response-locked CDA (−300 to 0 ms), as opposed to the stimulus-locked CDA reported in the main text, to ensure any observed differences in mean amplitude between conditions were not driven by differences in response-time (Ankaoua & Luria, 2023; Williams & Drew, 2021).

Experiment 1, Memory Set Size (MSS) 1–8

The response-locked CDA in Experiment 1 is shown in Figure 1. Similar to the stimulus-locked CDA, the response-locked CDA amplitude varied with MSS, $F(3, 57) = 2.81$, p = .047, $BF_{10} = .18$, and was larger for target present than target absent trials, $F(1, 19) = 21.31$, $p < .001$, $BF_{10} = 2.22e+6$. However, the MSS by target

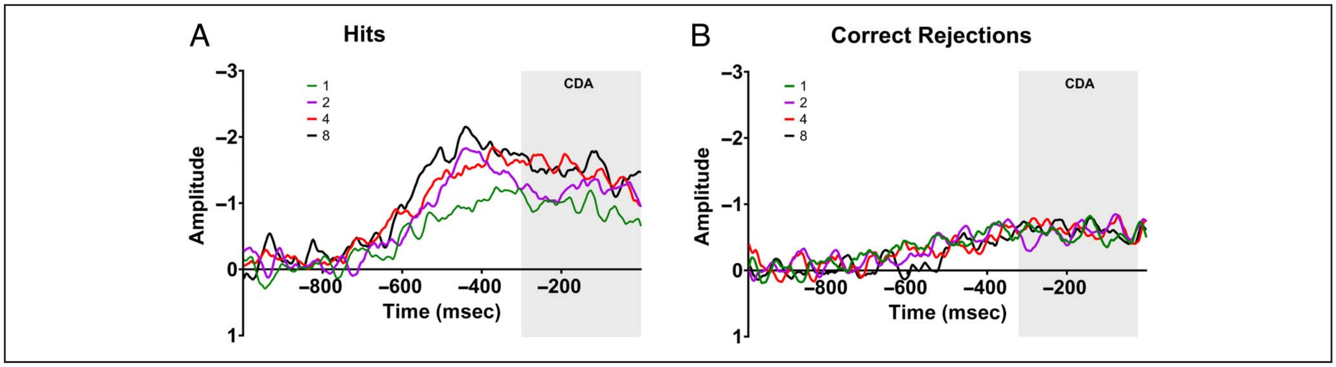

Figure A1. Response-locked lateralized waveforms for Experiment 1: (A) target present trials, (B) target absent trials.

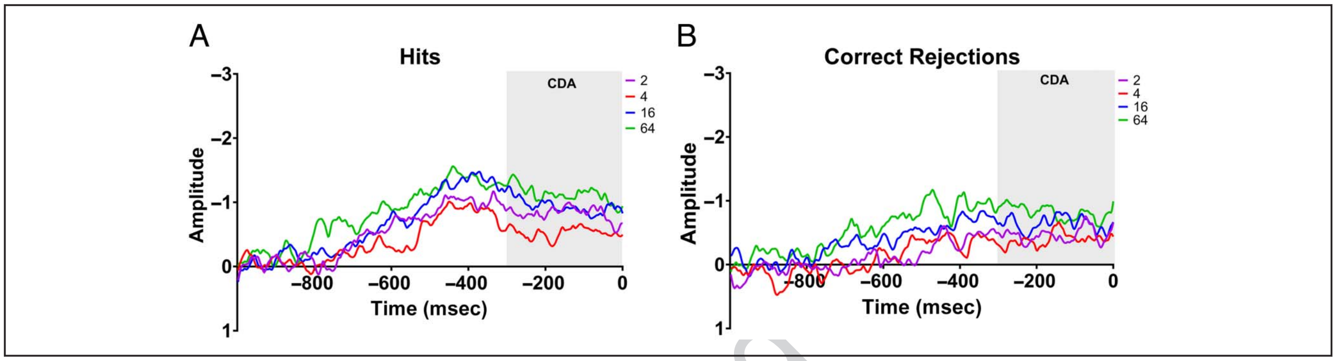

Figure A2. Response-locked lateralized waveforms for Experiment 2: (A) target present trials. (B) target absent trials.

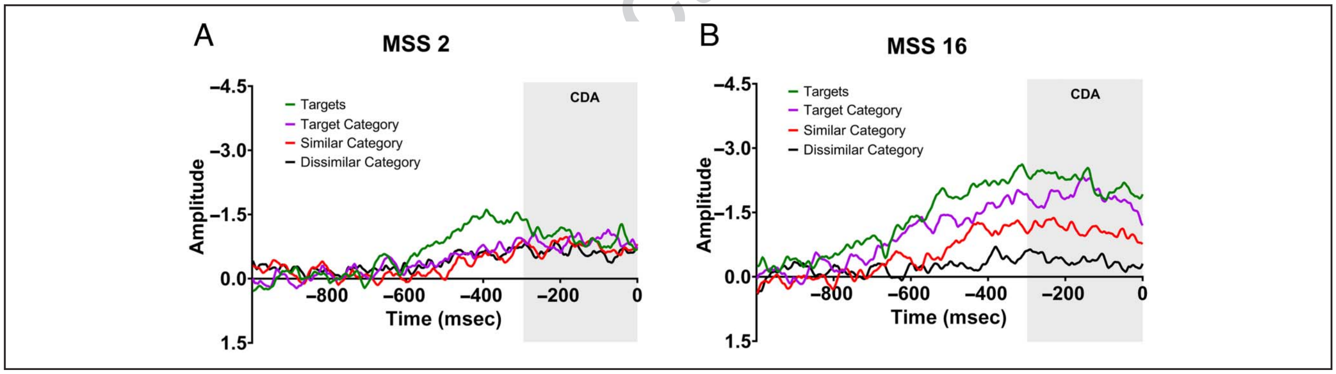

Figure A3. Response-locked waveforms for Experiment 3: (A) memory set size (MSS) 2. (B) MSS 16.

presence interaction was not statistically significant, $F(3, 57) = 1.2$, p = .32, $BF_{10} = .26$, and none of post-hoc tests survived the adjustments for multiple comparisons, all p-values > .05.

Experiment 2, MSS 2–64

The response-locked CDA in Experiment 2 is shown in Figure 2. Similar to the stimulus-locked CDA, the response-locked CDA amplitude also varied with MSS, $F(3, 81) = 4.18$, p = .008, $BF_{10} = .70$, so that the CDA amplitude was significantly smaller (less negative) for MSS 4 than for set size 64, p = .006. There were no significant differences between any of the other MSS comparisons, all $p > .05$. The main effect of target presence did not reach significance, $F(1, 27) = 3.69$, p = .07, $BF_{10} = 1.23$, and the set size by target presence interaction was not significant, $F(3, 81) = .12$, p = .95, $BF_{10} = .06$.

Experiment 2, Categorical Similarity, MSS 2 and 16

The response-locked CDA in Experiment 3 is shown in Figure 3. Similar to the stimulus-locked CDA, the amplitude of the response-locked CDA also increased with MSS, $F(1, 19) = 16.86$, p = .001, $BF_{10} = 52.81$, and varied with similarity between the attended item and the target category, $F(3, 57) = 10.26$, $p < .001$, $BF_{10} = 4070.98$. Similar to the stimulus-locked effects, the response-locked

CDA was larger for the targets than non-targets from the similar category, p = .02, and non-targets from the dissimilar category, $p < .001$, and larger for non-targets from the target category than non-targets from the dissimilar category, p = .001. None of the other comparisons were significant, all p-values > .05. Also the condition by MSS interaction was significant, $F(3, 57) = 7.15$, p = .002, $BF_{10} = 33.37$, reflecting that the CDA increased with MSS for targets, $p < .0001$, and non-targets from the target category, p = .0002, but not for non-targets from a similar, p = .18, and dissimilar category, p = .30.

Corresponding author: Iris Wiegand, Donders Institute for Brain, Cognition and Behaviour, Department of Neuropsychology and Rehabilitation Psychology, Radboud Universiteit, Thomas van Aquinostraat 4, Netherlands, or via e-mail: iris.wiegand@donders.ru.nl.

Data Availability Statement

The pre-processed data that entered the reported analyses are shared on the Open Science Framework. Experiment 1: https://osf.io/cdp4f/; Experiment 2: https://osf.io/dms4q/; Experiment 3: https://osf.io/zcsxv/.

Author Contributions

Lauren H. Williams: Conceptualization; Data curation; Formal analysis; Investigation; Methodology; Project administration; Writing—Original draft; Writing—Review & editing. Iris Wiegand: Conceptualization; Formal analysis; Methodology; Writing—Original draft; Writing—Review & editing. Mark Lavelle: Investigation; Methodology; Writing—Review & editing. Jeremy M. Wolfe: Conceptualization; Writing—Review & editing. Keisuke Fukuda: Writing—Review & editing. Marius V. Peelen: Writing—Review & editing. Trafton Drew: Conceptualization; Resources; Supervision; Writing—Review & editing.

Funding Information

This material is based upon work supported by the National Science Foundation Graduate Research Fellowship under Grant No. 1747505.

Diversity in Citation Practices

Retrospective analysis of the citations in every article published in this journal from 2010 to 2021 reveals a persistent pattern of gender imbalance: Although the proportions of authorship teams (categorized by estimated gender identification of first author/last author) publishing in the *Journal of Cognitive Neuroscience* (*JoCN*) during this period were M(an)/M = .407, W(oman)/M = .32, M/W = .115, and W/W = .159, the comparable proportions for the articles that these authorship teams cited were M/M = .549, W/M = .257, M/W = .109, and W/W = .085 (Postle and Fulvio, *JoCN*, 34:1, pp. 1–3). Consequently, *JoCN* encourages all authors to consider gender balance explicitly when selecting which articles to cite and gives them the opportunity to report their article's gender citation balance. The authors of this paper report its proportions of citations by gender category to be: M/M = .60; W/M = .32; M/W = .04; W/W = .04.

REFERENCES

Ankaoua, M., & Luria, R. (2023). One turn at a time: Behavioral and ERP evidence for two types of rotations in the classical mental rotation task. *Psychophysiology*, *60*, e14213. https://doi.org/10.1111/psyp.14213, PubMed: 36371685

Bays, P. M., Wu, E. Y., & Husain, M. (2011). Storage and binding of object features in visual working memory. *Neuropsychologia*, *49*, 1622–1631. https://doi.org/10.1016/j.neuropsychologia.2010.12.023, PubMed: 21172364

Boettcher, S. E. P., Drew, T., & Wolfe, J. M. (2018). Lost in the supermarket: Quantifying the cost of partitioning memory sets in hybrid search. *Memory & Cognition*, *46*, 43–57. https://doi.org/10.3758/s13421-017-0744-x, PubMed: 28786022

Brady, T. F., Konkle, T., Alvarez, G. A., & Oliva, A. (2008). Visual long-term memory has a massive storage capacity for object details. *Proceedings of the National Academy of Sciences, U.S.A.*, *105*, 14325–14329. https://doi.org/10.1073/pnas.0803390105, PubMed: 18787113

Brisson, B., & Jolicoeur, P. (2008). Express attentional re-engagement but delayed entry into consciousness following invalid spatial cues in visual search. *PLoS One*, *3*, e3967. https://doi.org/10.1371/journal.pone.0003967, PubMed: 19088847

Carlisle, N. B., Arita, J. T., Pardo, D., & Woodman, G. F. (2011). Attentional templates in visual working memory. *Journal of Neuroscience*, *31*, 9315–9322. https://doi.org/10.1523/JNEUROSCI.1097-11.2011, PubMed: 21697381

Cunningham, C. A., & Wolfe, J. M. (2014). The role of object categories in hybrid visual and memory search. *Journal of Experimental Psychology: General*, *143*, 1585–1599. https://doi.org/10.1037/a0036313, PubMed: 24661054

Curran, T. (2000). Brain potentials of recollection and familiarity. *Memory & Cognition*, *28*, 923–938. https://doi.org/10.3758/BF03209340, PubMed: 11105518

Curran, T., & Cleary, A. M. (2003). Using ERPs to dissociate recollection from familiarity in picture recognition. *Cognitive Brain Research*, *15*, 191–205. https://doi.org/10.1016/S0926-6410(02)00192-1, PubMed: 12429370

Delorme, A., & Makeig, S. (2004). EEGLAB: An open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of Neuroscience Methods*, *134*, 9–21. https://doi.org/10.1016/j.jneumeth.2003.10.009, PubMed: 15102499

Desimone, R., & Duncan, J. (1995). Neural mechanisms of selective visual attention. *Annual Review of Neuroscience*, *18*, 193–222. https://doi.org/10.1146/annurev.ne.18.030195.001205, PubMed: 7605061

Drew, T., Boettcher, S. E. P., & Wolfe, J. M. (2016). Searching while loaded: Visual working memory does not interfere with hybrid search efficiency but hybrid search uses working memory capacity. *Psychonomic Bulletin & Review*, *23*, 201–212. https://doi.org/10.3758/s13423-015-0874-8, PubMed: 26055755

Drew, T., Boettcher, S. E. P., & Wolfe, J. M. (2017). One visual search, many memory searches: An eye-tracking investigation of hybrid search. *Journal of Vision*, *17*, 5. https://doi.org/10.1167/17.11.5, PubMed: 28892812

Drew, T., Williams, L. H., Jones, C. M., & Luria, R. (2018). Neural processing of repeated search targets depends upon the stimuli: Real world stimuli engage semantic processing and recognition memory. *Frontiers in Human Neuroscience*, *12*, 460. https://doi.org/10.3389/fnhum.2018.00460, PubMed: 30519166

Eimer, M. (1996). The N2pc component as an indicator of attentional selectivity. *Electroencephalography and Clinical Neurophysiology*, *99*, 225–234. https://doi.org/10.1016/0013-4694(96)95711-9, PubMed: 8862112

Fukuda, K., Awh, E., & Vogel, E. K. (2010). Discrete capacity limits in visual working memory. *Current Opinion in Neurobiology*, *20*, 177–182. https://doi.org/10.1016/j.conb.2010.03.005, PubMed: 20362427

Grubert, A., Carlisle, N. B., & Eimer, M. (2016). The control of single-color and multiple-color visual search by attentional templates in working memory and in long-term memory. *Journal of Cognitive Neuroscience*, *28*, 1947–1963. https://doi.org/10.1162/jocn_a_01020, PubMed: 27458746

Grubert, A., & Eimer, M. (2016). All set, indeed! N2pc components reveal simultaneous attentional control settings for multiple target colors. *Journal of Experimental Psychology: Human Perception and Performance*, *42*, 1215–1230. https://doi.org/10.1037/xhp0000221, PubMed: 26950386

Guild, E. B., Cripps, J. M., Anderson, N. D., & Al-Aidroos, N. (2014). Recollection can support hybrid visual memory search. *Psychonomic Bulletin & Review*, *21*, 142–148. https://doi.org/10.3758/s13423-013-0483-3, PubMed: 23884688

Gunseli, E., Meeter, M., & Olivers, C. N. L. (2014). Is a search template an ordinary working memory? Comparing electrophysiological markers of working memory maintenance for visual search and recognition. *Neuropsychologia*, *60*, 29–38. https://doi.org/10.1016/j.neuropsychologia.2014.05.012, PubMed: 24878275

Hilimire, M. R., Mounts, J. R. W., Parks, N. A., & Corballis, P. M. (2011). Dynamics of target and distractor processing in visual search: Evidence from event-related brain potentials. *Neuroscience Letters*, *495*, 196–200. https://doi.org/10.1016/j.neulet.2011.03.064, PubMed: 21457759

Ikkai, A., McCollough, A. W., & Vogel, E. K. (2010). Contralateral delay activity provides a neural measure of the number of representations in visual working memory. *Journal of Neurophysiology*, *103*, 1963–1968. https://doi.org/10.1152/jn.00978.2009, PubMed: 20147415

Konkle, T., Brady, T. F., Alvarez, G. A., & Oliva, A. (2010). Conceptual distinctiveness supports detailed visual long-term memory for real-world objects. *Journal of Experimental Psychology: General*, *139*, 558–578. https://doi.org/10.1037/a0019165, PubMed: 20677899

Küper, K., Groh-Bordin, C., Zimmer, H. D., & Ecker, U. K. H. (2012). Electrophysiological correlates of exemplar-specific processes in implicit and explicit memory. *Cognitive, Affective, & Behavioral Neuroscience*, *12*, 52–64. https://doi.org/10.3758/s13415-011-0065-7, PubMed: 22038704

Küper, K., & Zimmer, H. D. (2018). The impact of perceptual changes to studied items on ERP correlates of familiarity and recollection is subject to hemispheric asymmetries. *Brain and Cognition*, *122*, 17–25. https://doi.org/10.1016/j.bandc.2018.01.006, PubMed: 29396208

Lavelle, M., Luria, R., & Drew, T. (2023). Incidental recognition reveals attentional tradeoffs shaped by categorical similarity. *Journal of Experimental Psychology: Human Perception and Performance*, *49*, 893–906. https://doi.org/10.1037/xhp0001128, PubMed: 37276126

Lopez-Calderon, J., & Luck, S. J. (2014). ERPLAB: An open-source toolbox for the analysis of event-related potentials. *Frontiers in Human Neuroscience*, *8*, 213. https://doi.org/10.3389/fnhum.2014.00213, PubMed: 24782741

Luck, S. J., & Hillyard, S. A. (1994). Electrophysiological correlates of feature analysis during visual search. *Psychophysiology*, *31*, 291–308. https://doi.org/10.1111/j.1469-8986.1994.tb02218.x, PubMed: 8008793

Luck, S. J., & Vogel, E. K. (1997). The capacity of visual working memory for features and conjunctions. *Nature*, *390*, 279–281. https://doi.org/10.1038/36846, PubMed: 9384378

Luria, R., Balaban, H., Awh, E., & Vogel, E. K. (2016). The contralateral delay activity as a neural measure of visual working memory. *Neuroscience & Biobehavioral Reviews*, *62*, 100–108. https://doi.org/10.1016/j.neubiorev.2016.01.003, PubMed: 26802451

Luria, R., & Vogel, E. K. (2011). Visual search demands dictate reliance on working memory storage. *Journal of Neuroscience*, *31*, 6199–6207. https://doi.org/10.1523/JNEUROSCI.6453-10.2011, PubMed: 21508243

Nako, R., Wu, R., Smith, T. J., & Eimer, M. (2014). Item and category-based attentional control during search for real-world objects: Can you find the pants among the pans? *Journal of Experimental Psychology: Human Perception and Performance*, *40*, 1283–1288. https://doi.org/10.1037/a0036885, PubMed: 24820441

Nartker, M. S., Alaoui-Soce, A., & Wolfe, J. M. (2020). Visual search errors are persistent in a laboratory analog of the incidental finding problem. *Cognitive Research: Principles and Implications*, *5*, 32. https://doi.org/10.1186/s41235-020-00235-4, PubMed: 32728864

Norman, K. A., Tepe, K., Nyhus, E., & Curran, T. (2008). Event-related potential correlates of interference effects on recognition memory. *Psychonomic Bulletin & Review*, *15*, 36–43. https://doi.org/10.3758/PBR.15.1.36, PubMed: 18605477

Nosofsky, R. M., Cox, G. E., Cao, R., & Shiffrin, R. M. (2014). An exemplar-familiarity model predicts short-term and long-term probe recognition across diverse forms of memory search. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *40*, 1524–1539. https://doi.org/10.1037/xlm0000015, PubMed: 24749963

Ort, E., & Olivers, C. N. L. (2020). The capacity of multiple-target search. *Visual Cognition*, *28*, 330–355. https://doi.org/10.1080/13506285.2020.1772430

Rugg, M. D., & Curran, T. (2007). Event-related potentials and recognition memory. *Trends in Cognitive Sciences*, *11*, 251–257. https://doi.org/10.1016/j.tics.2007.04.004

Shang, L., Yeh, L.-C., Zhao, Y., Wiegand, I., & Peelen, M. V. (2024). Category-based attention facilitates memory search. *eNeuro*, *11*, ENEURO.0012-24.2024. https://doi.org/10.1523/ENEURO.0012-24.2024, PubMed: 38331577

Shiffrin, R. M., & Steyvers, M. (1997). A model for recognition memory: REM-retrieving effectively from memory. *Psychonomic Bulletin & Review*, *4*, 145–166. https://doi.org/10.3758/BF03209391, PubMed: 21331823

Turoman, N., Tivadar, R. I., Retsa, C., Murray, M. M., & Matusz, P. J. (2021). Towards understanding how we pay attention in naturalistic visual search settings. *Neuroimage*, *244*, 118556. https://doi.org/10.1016/j.neuroimage.2021.118556, PubMed: 34492292

Vogel, E. K., & Machizawa, M. G. (2004). Neural activity predicts individual differences in visual working memory capacity. *Nature*, *428*, 748–751. https://doi.org/10.1038/nature02447, PubMed: 15085132

Wiegand, I., Finke, K., Müller, H. J., & Töllner, T. (2013). Event-related potentials dissociate perceptual from response-related age effects in visual search. *Neurobiology of*

Aging, *34*, 973–985. https://doi.org/10.1016/j.neurobiolaging.2012.08.002, PubMed: 22921866

Wiegand, I., Napiórkowski, N., Töllner, T., Petersen, A., Habekost, T., Müller, H. J., et al. (2018). Event-related electroencephalographic lateralizations mark individual differences in spatial and nonspatial visual selection. *Journal of Cognitive Neuroscience*, *30*, 482–497. https://doi.org/10.1162/jocn_a_01221, PubMed: 29244636

Wiegand, I., Töllner, T., Habekost, T., Dyrholm, M., Müller, H. J., & Finke, K. (2014). Distinct neural markers of TVA-based visual processing speed and short-term storage capacity parameters. *Cerebral Cortex*, *24*, 1967–1978. https://doi.org/10.1093/cercor/bht071, PubMed: 23535180

Wiegand, I., & Wolfe, J. M. (2020). Age doesn't matter much: Hybrid visual and memory search is preserved in older adults. *Aging, Neuropsychology, and Cognition*, *27*, 220–253. https://doi.org/10.1080/13825585.2019.1604941, PubMed: 31050319

Wiegand, I., Wolfe, J. M., Maes, J. H. R., & Kessels, R. P. C. (2024). Incidental learning of temporal and spatial associations in hybrid search. *Visual Cognition*, 1–16. https://doi.org/10.1080/13506285.2024.2346991

Williams, L. H., & Drew, T. (2021). Maintaining rejected distractors in working memory during visual search depends on search stimuli: Evidence from contralateral delay activity. *Attention, Perception, & Psychophysics*, *83*, 67–84. https://doi.org/10.3758/s13414-020-02127-7, PubMed: 33000442

Wolfe, J. M. (2012). Saved by a log: How do humans perform hybrid visual and memory search? *Psychological Science*, *23*, 698–703. https://doi.org/10.1177/0956797612443968, PubMed: 22623508

Wolfe, J. M. (2021). Guided Search 6.0: An updated model of visual search. *Psychonomic Bulletin & Review*, *28*, 1060–1092. https://doi.org/10.3758/s13423-020-01859-9, PubMed: 33547630

Wolfe, J. M., Alaoui Soce, A., & Schill, H. M. (2017). How did I miss that? Developing mixed hybrid visual search as a 'model system' for incidental finding errors in radiology. *Cognitive Research: Principles and Implications*, *2*, 35. https://doi.org/10.1186/s41235-017-0072-5, PubMed: 28890920

Wolfe, J. M., Boettcher, S. E. P., Josephs, E. L., Cunningham, C. A., & Drew, T. (2015). You look familiar, but I don't care: Lure rejection in hybrid visual and memory search is not based on familiarity. *Journal of Experimental Psychology: Human Perception and Performance*, *41*, 1576–1587. https://doi.org/10.1037/xhp0000096, PubMed: 26191615

Wu, R., Scerif, G., Aslin, R. N., Smith, T. J., Nako, R., & Eimer, M. (2013). Searching for something familiar or novel: Top–down attentional selection of specific items or object categories. *Journal of Cognitive Neuroscience*, *25*, 719–729. https://doi.org/10.1162/jocn_a_00352, PubMed: 23281777

Yonelinas, A. P. (2002). The nature of recollection and familiarity: A review of 30 years of research. *Journal of Memory and Language*, *46*, 441–517. https://doi.org/10.1006/jmla.2002.2864

Zimmer, H. D., & Ecker, U. K. H. (2010). Remembering perceptual features unequally bound in object and episodic tokens: Neural mechanisms and their electrophysiological correlates. *Neuroscience & Biobehavioral Reviews*, *34*, 1066–1079. https://doi.org/10.1016/j.neubiorev.2010.01.014, PubMed: 20138910