



Long-term memory and working memory compete and cooperate to guide attention

Zall Hirschstein¹ · Mariam Aly¹

Accepted: 2 October 2022
© The Psychonomic Society, Inc. 2022

Abstract

Multiple types of memory guide attention: Both long-term memory (LTM) and working memory (WM) effectively guide visual search. Furthermore, both types of memories can capture attention automatically, even when detrimental to performance. It is less clear, however, how LTM and WM cooperate or compete to guide attention in the same task. In a series of behavioral experiments, we show that LTM and WM reliably cooperate to guide attention: Visual search is faster when both memories cue attention to the same spatial location (relative to when only one memory can guide attention). LTM and WM competed to guide attention in more limited circumstances: Competition only occurred when these memories were in different dimensions – particularly when participants searched for a shape and held an accessory color in mind. Finally, we found no evidence for asymmetry in either cooperation or competition: There was no evidence that WM helped (or hindered) LTM-guided search more than the other way around. This lack of asymmetry was found despite differences in LTM-guided and WM-guided search overall, and differences in how two LTMs and two WMs compete or cooperate with each other to guide attention. This work suggests that, even if only one memory is currently task-relevant, WM and LTM can cooperate to guide attention; they can also compete when distracting features are salient enough. This work elucidates interactions between WM and LTM during attentional guidance, adding to the literature on costs and benefits to attention from multiple active memories.

Keywords Attention: Interactions with memory · Memory: Long-term memory · Visual search

Introduction

Research on attention commonly distinguishes *top-down* from *bottom-up* attention. Bottom-up attention is studied by characterizing how salient information in the external world captures attention. Top-down attention is studied by manipulating

how plans and goals affect what is attended, typically with explicit instructions. Although useful, this dichotomy omits another key conceptual distinction: That both of these forms of attention can be influenced by different types of memories (Aly & Turk-Browne, 2017; Awh et al., 2012; Chen & Hutchinson, 2019; Hutchinson & Turk-Browne, 2012). These memories may derive from distinct sources, perhaps recently acquired and maintained in working memory (WM) or stored in long-term memory (LTM) (Nobre & Stokes, 2019). It can be argued, in fact, that this prospective property of memory – its ability to guide attention – effectively constitutes its ecological purpose (Nobre & Stokes, 2019; van Ede & Nobre, 2023).

Despite recognition that both WM and LTM guide attention, it is largely unknown how they may interact when given the opportunity to guide attention together. In fact, there is relatively little research on how multiple active memories may guide attention, nor consensus regarding the conditions under which guidance from multiple items is exhibited (see, e.g., Frätescu et al., 2019, for a relevant debate). Even less well studied is how memories of different types (e.g., LTM and WM) may cooperate or compete with one another to

Significance At any given moment, multiple memories may be in mind. This can include information we learned some time ago (long-term memory; LTM) and information we recently acquired (working memory; WM). How do such memories interact to direct how we pay attention? We found that LTM and WM cooperate to guide attention, leading to better performance when both memories (vs. only one) suggest a common goal. LTM and WM sometimes competed, however, leading to worse performance when they suggested different goals. These results therefore highlight the dynamic ways that multiple memories can interact to guide our attention.

✉ Zall Hirschstein
zall.hirschstein@gmail.com

¹ Department of Psychology, Columbia University, 406 Schermerhorn Hall, 1190 Amsterdam Avenue, MC 5501, New York, NY 10027, USA

guide attention when active in the same task. Some studies have examined attentional guidance from multiple co-active features from either WM or from LTM (Bahle et al., 2020; Chen & Du, 2017; Zhang et al., 2018; Zhou et al., 2020), but don't examine LTM-WM interactions. These findings indicate, however, that at least two WM or two LTM items can be co-active in memory-guided attention, and that they are capable of both cooperating and competing for attentional resources. The limited research on interactions between LTM and WM during attentional guidance has found that they can both guide attention in the same task, but this work tends to examine LTM-WM interactions in only one direction (e.g., how WM may interfere with LTM-guided search; Günseli et al., 2016) or lacks baselines to permit distinction between cooperation and competition (e.g., Schwark et al., 2013).

The present study aims to fill these gaps by exploring how LTM and WM might be used to cooperatively guide attention together, but also how they may compete – slowing attentional guidance when the two representations conflict as to where attention should be directed. Addressing these questions allows us to propose a model of how memory-guided attention is orchestrated when multiple memories are active in a given task.

We were inspired by work showing that both WM and LTM can guide attention (Baddeley, 2003; Desimone & Duncan, 1995; Fan & Turk-Browne, 2016; Stokes et al., 2012; Summerfield et al., 2006). When information in WM and LTM is *consistent* — i.e., suggests the same attentional goal — we might therefore expect cooperation between these memories. This would lead to behavioral facilitation relative to when only one representation guides attention. Such cooperative facilitation between memory types would accord with the literature on “redundancy gains” whereby a behavioral advantage is conferred when multiple pieces of information support the same decision, relative to when only one supports it (Bahle et al., 2020; Danek & Mordkoff, 2011; Fan & Turk-Browne, 2016; Miller & Low, 2001). That is, an environmental stimulus matching two active memory representations — regardless of whether sourced from WM or LTM — should more efficiently guide attention than a stimulus that matches a WM or LTM in isolation. To our knowledge, no evidence suggests that such guidance should be asymmetrical, in that WM would help LTM-guided search more or less than LTM would help WM-guided search; so, in the case of cooperation, we expect dual guidance from WM and LTM will entail symmetrical facilitation.

It is not always the case, however, that memories suggest the same attentional goal. In such a scenario, how might WM and LTM compete for attention? Research has demonstrated that attentional guidance from WM or LTM can occur automatically, even when irrelevant to the current task and employing them would lead to distraction from current attentional goals (Downing, 2000; Fan & Turk-Browne, 2016;

Günseli et al., 2016; Nickel et al., 2020; Soto et al., 2005, 2008). This suggests that when WM and LTM are placed in competition with one another — that is, when a WM representation and an LTM representation suggest attentional goals that are *inconsistent* with one another, even if only one memory is currently task-relevant — behavioral deficits should emerge relative to when only one memory guides attention.

We test two alternative hypotheses about how WM and LTM may compete with one another. First, because both WM and LTM can lead to automatic attentional capture (Downing, 2000; Fan & Turk-Browne, 2016; Günseli et al., 2016; Nickel et al., 2020; Soto et al., 2005, 2008), one might hypothesize that competition would be symmetrical. That is, irrelevant WM representations may hinder LTM-guided search just as much as the other way around. This hypothesis would be consonant with findings showing common representational formats for WM and LTM (Cowan, 1993; Cowan, 1998; Cowan, 2008; Fuster, 1997; Ranganath & Blumenfeld, 2005; Vo et al., 2022), and/or the notion that activated representations in LTM must be “placed” into WM (Atkinson & Shiffrin, 1968). If this is the case, then there might be no functional difference between items “in WM” and (activated) items “in LTM” — both should be represented similarly. If so, they should compete with one another symmetrically, as noted above.

An alternative hypothesis, however, is that competition between WM and LTM may be asymmetrical. Research has suggested that a “flexible gate” governs the transmission of information from LTM to WM: LTM is admissible only when it would be useful given current task demands, and blocked if it would hinder performance (Mızrak & Oberauer, 2021; Oberauer et al., 2017; see also Verschooren et al., 2021). If so, there should be a larger cost to performance when using LTM to guide attention in the face of distraction from WM, relative to when WM is used in the face of distraction from LTM. That is, WM-guided attention may be protected from interference by accessory LTM representations when they would hurt performance; but LTM-guided attention may be more prone to distraction from irrelevant WM representations.

Mechanistically, this asymmetry may be explained by the concentric activation model (Oberauer, 2002, 2009), which distinguishes between three functional states of WM: The activated part of LTM (aLTM), the region of direct access (RDA), and a single-item focus of attention (FoA). According to this model, the memory selected for the next cognitive operation should be provided the privileged status afforded to items in the FoA. WM information that is accessory to the primary task at hand — currently irrelevant but still active — should be represented in the RDA. It should therefore be distinguishable from accessory LTM information, theoretically represented in aLTM. If these unique functional states confer differential effects when guiding attention, we may observe competitive asymmetry: Items in the RDA may be

more likely to capture attention than items in aLTM. Asymmetry in competition may also arise because of differences in susceptibility to interference. WM is more vulnerable to interference, and requires active maintenance, or it may be lost; LTM on the other hand, if temporarily lost during WM-guided attentional processes, may still be capable of being reactivated at a later time. Given this, irrelevant WMs may be more distracting than irrelevant LTMs.

To test these hypotheses, we use a novel paradigm that allows us to dissociate cooperative and competitive interactions between WM and LTM during memory-guided search. Participants performed a task in which they searched for an item maintained in WM — or activated from LTM — while holding in mind an accessory item cued from LTM or WM. Three search display conditions were used: The accessory item either guided attention to the same spatial location as the prioritized memory, a different spatial location, or was not present in the display. Following each trial, participants precisely reported features of the accessory item. This paradigm allowed us to determine how WM and LTM compete or cooperate to guide attention.

Experiment 1

Methods

Participants

Data were collected until the final sample comprised 115 participants who met the inclusion criteria (described below). The final sample size of 115 was determined *a priori* based on a review of relevant literature, in which RT was a dependent variable and memory contents competed for attentional control (e.g., Beck et al., 2012 (807 observations/condition); Chen & Du, 2017 (across-experiment average: 1,448 observations/condition); Fan & Turk-Browne, 2016 (across-experiment average: 1,380 observations/condition); van Moorselaar et al., 2014 (across-experiment average: 1,020 observations/condition); Zhang et al., 2018 (1,152 observations/condition)). The total number of observations per condition (trials \times participants) were calculated for these relevant studies (Baker et al., 2021) and the sample size for this study determined by approximately matching the highest values from that calculation while accommodating specific constraints of our study (e.g., need to balance the number of times a given color was prompted). Given these constraints, we obtained 1,725 observations/condition (90 trials \times 115 participants/six conditions).

To meet that target sample size, 320 participants were recruited for an online study using Prolific (www.prolific.co; 251 participants) or the Columbia University Psychology Department Participant Pool (69 participants). Those

recruited on Prolific were pre-screened for English fluency, nationality (only USA), and age (18–40 years). All participants provided informed consent to a protocol approved by the Columbia University Institutional Review Board, and received \$6.50/h or course credit, respectively, as compensation. 137 of these participants were unable to pass one of the Testing Phases; their participation was terminated early and they are not included in the final sample.

To ensure that we had adequate data across all six conditions for each participant, and that each participant was doing the task as instructed, we set an *a priori* threshold to include only those participants who responded correctly to both components of a given trial (as described in *Procedure/Search Phase*) in $\geq 50\%$ of trials. Sixty-two participants from the remaining sample did not reach this criterion, and thus none of their data are included in the final sample. Following these rejections, the final sample comprised 115 participants, as noted above ($M_{\text{age}} = 26.2 \pm 6.6$ years, $M_{\text{education}} = 14.9 \pm 2.4$ years). Sixty-one of these participants identified as women and six as non-binary; the rest identified as men. Of this final sample, 71.4% identified as White, 19.6% as Asian, 9.8% as Black or African American, 4.8% as American Indian/Alaskan Native, 0.9% as Native Hawaiian or Pacific Islander, and 2.7% identified as part of a different racial group; in addition, 11.6% of these participants identified their ethnicity as Hispanic or Latino.

Participants were not screened for color blindness. However, individuals were only included in the final sample if they were able to perform well on the Training and Test Phases and responded correctly to both components of a Search Phase trial on $\geq 50\%$ of trials. Is therefore unlikely that included participants were color blind, or, if they did have some degree of color blindness, it did not interfere with their ability to perform well on the task.

Stimuli

Color generation Two sets of five colors were generated in Hue, Saturation and Luminance (HSL) colorspace (see Fig. 1a). Saturation and luminance were held constant — at 100 and 50, respectively — so that colors varied only along the hue dimension. Colors were not entirely equidistant from one another on the HSL colorwheel due to subjective similarity of the green hues. A constant “central green” hue (115° , between the upper and lower bounds of subjective similarity: 75° and 156°) was always present in the first of the two sets. The rest of the hue wheel was then equally divided (each hue was 31° apart) to generate the remaining nine colors present in the two sets. Colors alternated between sets, so that no set contained colors within 62° of each other. A variable buffer (between 0° and 31°) was added at the “starting point” of the hue generation process to allow the generated sets to differ from one another across blocks. The colors were generated in sets (as

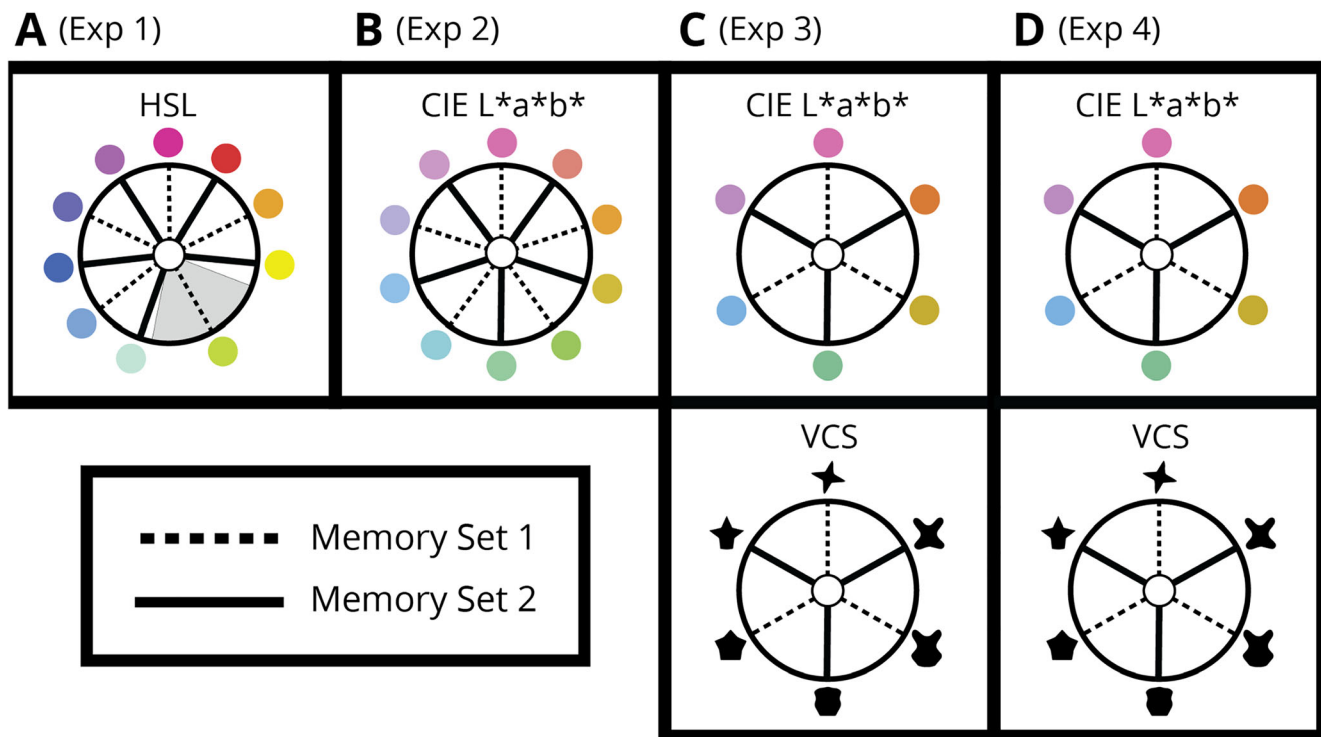


Fig. 1 Example generation of color sets (**a–d**) and shape sets (**c–d**) for each experiment. For more details, see the Color Generation or Shape Generation section of each experiment's methods. **a** In Experiment 1, ten colors were generated on a hue colorwheel in HSL colorspace. The gray slice of the colorwheel denotes the buffer zone around the “center green” hue. All other colors were separated from one other by 31° on the hue colorwheel. **b** In Experiment 2, ten colors were generated on the CIE

L*a*b* colorwheel at equidistant intervals (36°). **c** In Experiment 3, six colors were generated on the CIE L*a*b* colorwheel at equidistant intervals (60°). Likewise, six shapes were generated along the VCS shapewheel at equidistant intervals (60°). **d** In Experiment 4, the color and shape generation matched that of Experiment 3. Bold lines indicate colors (or shapes) in one set; dashed lines indicate colors (or shapes) in the other set

opposed to random selection from the color wheel) to be consistent with prior literature in the attentional capture field (Anderson & Halpern, 2017; Nickel et al., 2020) as well as to ease memorization in the Training Phase.

Color assignment After being generated, we randomly assigned color sets to memory conditions. The first set of colors was assigned to the WM color set and the second set of colors was assigned to the LTM color set half of the time; the assignment of color sets to memory conditions was reversed the other half of the time to ensure that any given color was equally likely to be present in the WM or LTM color set. Two colors were cued on each trial: One LTM color (that would be retrieved from LTM) and one WM color (that would be presented on the screen). The two colors on any given trial were required to be >90° apart on the hue colorwheel so as to be visually distinct; this constraint meant that each color was paired with three colors in the other set, permitting the use of 15 color pairs. There were 30 trials per block, so each color pair was cued twice per block. These color pairs were assigned to conditions such that each color was cued six times per block, once in each of the six conditions (described below). Therefore, each color was prompted (i.e., searched for) three

times and unprompted (i.e., not searched for, but reported after search) three times.

Colorwheel The HSL colorwheel used in Experiment 1 was adapted from project iro.js (github.com/jaames/iro.js) and modified to respond to user input and fit the needs of this experiment. First, to match the generated colors (see *Stimuli/Color generation*), saturation was held constant at 100 and luminance at 50. Thus, the colorwheel allowed just the selection of hues along a 360° circle. Participants could hover their cursor over the colorwheel to change the color of a central circle placed inside of the colorwheel to allow participants to view the color fully. Clicking selected the color corresponding to the clicked location on the hue colorwheel. At each presentation of the colorwheel screen, the colorwheel was rotated by 0–359°. This was done so that participants had to attend to, and encode, color itself rather than a relative location on the screen.

Scenes Ninety color scene images were manually selected from the “Massive Memory” Scene Categories database (Konkle et al., 2010). We avoided using multiple scene images from categories

that had substantial visual overlap (e.g., only one image was selected from the two categories “inside car” and “inside bus”), and avoided scene images that contained text. All images had a resolution of 256×256 . Scene images were randomized into three groups – one for each block of the experiment – the order of which was randomly determined per participant. Additionally, the order of presentation of these images was randomized within each block.

Procedure

Upon recruitment, participants were directed via URL to the experiment, which was hosted on Gorilla (www.gorilla.sc). Instructions were first provided describing the structure of the experiment: There were three blocks of the experiment, and each block comprised three successive phases: A Training Phase, a Testing Phase, and a Search Phase (see Fig. 2). Participants were informed that breaks were provided between blocks of the experiment but that breaks were not permitted within the blocks. (Each instruction page had a maximum allowable time for viewing, after which the instructions would advance if the participant had not manually done so to ensure that participants could not take long breaks within blocks.) Lastly, participants were asked to use a computer mouse and to turn the brightness on their screen to maximum for the duration of the experiment (to facilitate color discrimination), though adherence to this could not be properly assessed due to the online nature of the experiment.

At the beginning of each block, two sets of five colors were generated (see *Methods/Stimuli/Color generation* for more details). As noted above, one of these sets was assigned as the LTM set (to be learned in the Training Phase and later searched for in the Search Phase) and the other was assigned as the WM set (to be searched for in the Search Phase).

Training Phase Each Training Phase (Fig. 2a) consisted of 30 unique scene-color pairs presented adjacent to one another on the screen for 3,000 ms. Each of the five LTM colors was paired with six different scenes. The colors were presented as filled circles. Whether the scene was presented on the left or the right side of the screen (and, concomitantly, whether the color was on the right or left; for the Search Phase as well) was randomly chosen at the beginning of the first block and then switched between each subsequent block.

Participants were asked to memorize the association between each image and its color. They were informed that they would later be tested on the scene-color pairs (Testing Phase) by reporting the color associated with each scene. After the presentation of a scene-color pair for 3,000 ms, a colorwheel was presented for $\leq 8,000$ ms. Participants were tasked with selecting the color that they had viewed on the previous screen, to facilitate encoding. Upon submission of the color and provision of feedback about their accuracy (see below),

the participant could manually advance the screen before the 8,000 ms expired.

Upon submission of reports on the colorwheel, participants received visual feedback about how close their report was to the correct color: A dotted line, rendered in the color reported, was drawn from the center of the colorwheel to the reported color; and – to give participants an estimate of how close they were – a solid line, rendered in the correct color, was drawn from the center to the correct color. In addition to this visual feedback, textual feedback was provided above the colorwheel. During the Training Phase, a report 0° from the correct color triggered the textual feedback, “*Perfect! Exactly 0 degrees off*”; $\leq 4^\circ$, “*Fantastic! Only X degrees off*”; $\leq 10^\circ$, “*Close! Just X degrees off*”; $\leq 30^\circ$, “*Incorrect. X degrees off*”; $> 30^\circ$, “*Paying attention? X degrees off*.” (The “X” in each of these feedback statements was replaced with the number of degrees a participant’s report was from the correct color.) On the colorwheel used in the Search Phase, however – on the unprompted report screen (see *Methods/Procedure/Search Phase*) – only the dotted line described above was used. That is, on the Search Phase colorwheel, no indication was given as to the correct color when one was selected.

Each of the 30 scene-color pairs was repeated three times in the Training Phase. All scene-color pairs were presented once before a given scene-color pair was repeated. A brief (≤ 33 s) break was provided after one cycle of the 30 scene-color presentations, during which participants were informed that the pairs would now repeat, alongside an encouraging message.

Testing Phase During the Testing Phase, participants were told to report the color associated with any given scene when presented with just the scene. They were informed that if they could not accurately report a color after eight presentations of a given scene, the experiment would terminate early.

The Testing Phase (Fig. 2b) consisted of a centrally presented scene (3,000 ms) that had been associated with a color in the Training Phase. Following this, the colorwheel was presented centrally and the participants reported the color previously associated with that scene ($\leq 8,000$ ms).

After the color report, participants saw the same visual feedback regarding their accuracy as they did in the Training Phase. The textual feedback, however, was altered such that a report $\leq 10^\circ$ from the correct color triggered the feedback “*Correct*,” while a report of $> 10^\circ$ away triggered “*Incorrect*.” In accordance with this textual feedback, participants’ reports were considered “correct” if they were $\leq 10^\circ$ from the correct color on the colorwheel. If a scene was responded to incorrectly, it was added to the next cycle of the Testing Phase. The second cycle began after the first cycle through all 30 scenes, and subsequent cycles through the incorrectly responded-to scenes continued until participants were able to respond correctly to each scene.

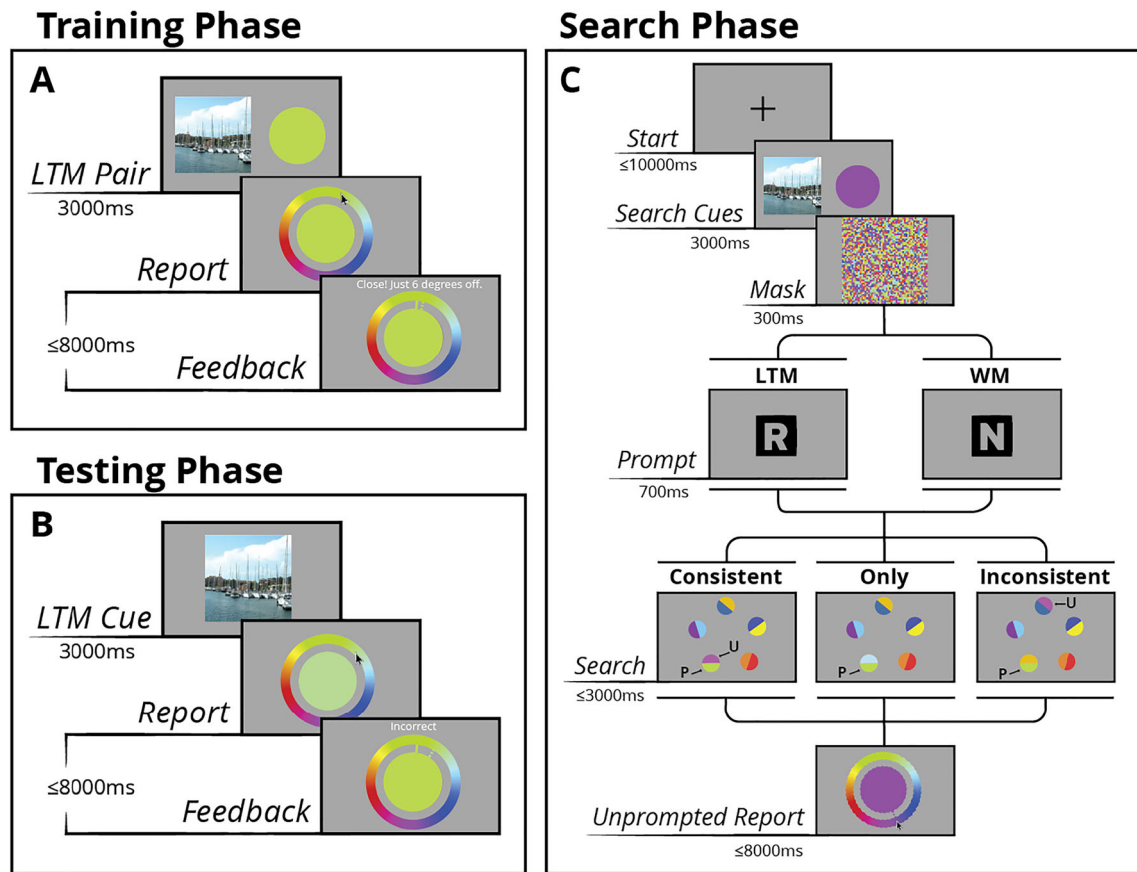


Fig. 2 Experiment 1 procedure. Participants completed three blocks, each consisting of a Training Phase, a Testing Phase, and a Search Phase. **a** The Training Phase of each block consisted of 30 scene-color pairs to be encoded into LTM. Participants first viewed each scene and its associated color, then immediately reported the color associated with each scene on the hue colorwheel. Each scene-color pair was presented three times during this phase. **b** The Testing Phase required that participants recall and accurately report the color associated with each scene, when presented with the scene alone. **c** The Search Phase presented two cues on each trial: An image that had been encoded into LTM, for which participants had to remember the associated (LTM) color; and a "new" (WM) color in a filled circle. After being prompted to search for either the "remembered"

(R) or "new" (N) color, participants clicked the prompted color in the search display. There were three conditions for this search display: *Consistent* trials contained both the prompted and unprompted colors in a single circle; *Only* trials contained the prompted color only (and not the unprompted color); and *Inconsistent* trials contained both the prompted and unprompted colors, but these two colors were placed inside separate circles on the display. After the search display, participants were asked to report the unprompted color on the hue colorwheel. The example search and unprompted color report displays in this figure are for a trial in which the LTM ("remembered") color is prompted. For illustration purposes, the example search displays are labeled to denote the location of the prompted ("P") and unprompted ("U") colors

As they had been instructed, if a participant could not respond correctly to a scene after eight test cycles, the experiment terminated early. Between each test cycle, participants were provided with a brief (≤ 33 s) break screen that reported the percentage of scenes they had responded to correctly so far, and reminded them of the need to be able to respond to all scenes correctly, alongside an encouraging message.

Search Phase After successful completion of the Testing Phase, participants read instructions for the Search Phase, which consisted of a visual search task and a post-search color memory task. After passing a brief quiz to ensure they had understood the Search Phase procedure, participants began the Search Phase.

In the Search Phase (Fig. 2c), participants were first presented with a central fixation cross (for ≤ 10 s). To begin each search trial, they clicked the middle of the cross. This was done to center their mouse cursor relative to the browser window when each trial started. Participants received a reminder to "click the center of the cross to start the trial" after 7 s, and the trial would begin if three more seconds elapsed without a response. The following screen contained a scene image that had been previously associated with a color in the Training Phase. Alongside this scene, a colored circle was presented, like in the Training Phase; this circle was, however, filled with a color from the WM color set. This screen was presented for 3,000 ms. At this point in the Search Phase, participants had been instructed to "simultaneously" keep in mind this "new" (WM) color and bring to mind the "remembered" (LTM)

color previously associated with the scene image during the Training Phase. Participants were informed that both colors would be equally important.

A brief (300 ms) color mask was then flashed on the screen to flush sensory information. This mask was 400×400 pixels, and for every search trial each pixel was filled with a color randomly selected from HSL space (only selected along the hue dimension, otherwise with the same parameters as the color wheel; i.e., saturation was held constant at 100 and luminance at 50).

Participants were then prompted to search for one or the other of the two colors. A black letter “N” presented in the center of the screen prompted the participant to search for the “new” color (WM prompt); and an “R” prompted the participant to search for the “remembered” color (LTM prompt). This prompt was displayed on the screen for 700 ms. Each trial, therefore, had both a prompted color and an unprompted color.

The search display was then presented for $\leq 3,000$ ms. The search display contained five circles at equal eccentricity from the center of the browser window and equidistant from one another. These circles were rotated around the center of the browser window in each search trial at a random angle between 0° and 359° . Each circle was divided in two, yielding ten differently colored half-circles. Each of the five circles was also rotated around its own central axis (the dividing line between the two colors) by a randomly selected angle across trials. These angles were selected to be equidistant from one another (e.g., for the five circles, each circle was rotated by 4° , 76° , 148° , 220° , or 292°). The purpose of this rotation was to potentially reduce participants’ ability to predict the location of stimuli on a given search display, in an attempt to more heavily tax visual search processes.

Participants were tasked with clicking on the half-circle that contained the prompted color as fast as possible. There were six conditions in total: As noted above, participants could be prompted to search for either the WM color or the LTM color. Within these two conditions, a search trial could be Consistent, Only, or Inconsistent. *Consistent* trials occurred when one of the five circles contained both the prompted and the unprompted color; thus, both the LTM color and the WM color guided attention to a spatially consistent location. *Only* trials occurred when the unprompted color was not in the display at all, and only the prompted color was in the display. *Inconsistent* trials occurred when both the prompted and the unprompted colors were in the display, but were in separate circles; thus, WM and LTM would guide attention to different (inconsistent) locations. The location of the prompted color was randomly determined on each trial. The location of the unprompted color was necessarily the half-circle adjacent to the prompted color on Consistent trials. On Inconsistent trials, the location of the unprompted color was randomly determined, following the constraint that it could not be the half-

circle adjacent to the prompted color. The unprompted color was not displayed on Only trials.

The remaining half-circles were filled with distractor colors. For each search display, eight or nine “distractor” colors (depending on condition) were generated in the same colorspace. These distractor colors were generated at random on the hue colorwheel, but at a minimum distance of 30° from both the LTM and WM colors cued on that trial, and at a minimum distance of 20° from other distractors. No circle could contain two colors that were $<45^\circ$ apart. Additionally, no distractor could be drawn within the “center green” buffer zone described above (under Stimuli/Color Generation), as these potential distractors were deemed to be too indistinguishable from the “center green” color itself. Upon clicking on any half-circle, the experiment advanced to the next screen.

After responding on the search display, participants were once again presented with the colorwheel. On this colorwheel, participants were tasked with reporting the *unprompted* color. That is, they were told to report the color that was cued at the beginning of the trial – either directly with a color (WM color) or indirectly with a scene (LTM color) – but was *not* searched for in the search display. We refer to this as the unprompted color report. Thus, if a given search trial was LTM-prompted, then participants would report the WM color associated with this trial and vice versa. Unlike in the Training and Testing Phases, participants were not provided with any feedback about their precision on this colorwheel. Participants completed one search trial for each scene-color combination for a total of 30 search trials per block (90 search trials total); thus, each scene was shown once. Half of the trials were LTM-prompted, the other half WM-prompted; one third of the trials were Consistent, one third Only, and one-third Inconsistent (divided equally between WM and LTM conditions).

At the conclusion of each block, participants were given a maximum 5-min break before advancing to the next block. At the start of each phase after the first block, participants were offered the opportunity to re-read or to skip the instructions for that phase. Upon successful completion of all three blocks, or unsuccessful completion of a Testing Phase, participants reported their demographic information and completed a brief debriefing questionnaire before being returned to their recruitment website and compensated.

Data

Exclusions Only those trials in which participants responded correctly to both the search display and the unprompted color report were included. These trials are referred to here as “both-correct.” This criterion was implemented to ensure that the trials in the main analyses were those for which participants maintained representations of both the prompted and the unprompted color. Search trials were considered correct if participants clicked within the half-circle filled with the prompted

color while the search display was on the screen. The unprompted color report was considered correct if participants' report was $\leq 20^\circ$ from the correct hue. This *a priori* threshold was selected to remain at least 10° away from the nearest possible distractor color, while allowing more error than the Testing Phase to accommodate less precision in reporting WM colors (participants had not practiced reporting colors from the WM color set until the Search Phase). Upon visualization of participants' responses, this threshold appeared to capture roughly 80% of the response distribution. An average of 67% of Search Phase trials were both-correct.

Analysis We modeled the primary dependent variable, log-transformed response time (RT) on the search task, with multilevel linear regression on participants' trial-wise data.

This analysis was conducted using R, version 4.0.5 (R Core Team, 2022), with the `lmer` function from the `lme4` package (Bates et al., 2015).

To test whether cooperation and competition arise between WM and LTM representations in this paradigm, we fit a mixed-effects model on log RT that included prompt (LTM or WM) and search display condition (Consistent, Only, or Inconsistent) as fixed factors, along with their interaction, and included random effects for the intercept and prompt terms to allow those estimates to vary between participants. That is: `lmer(log(RT) ~ prompt*searchCondition + (1 + prompt | Participant))`. Trial prompt was effect-coded (LTM = 1, WM = -1) to allow the other coefficients to be estimated at the grand mean of LTM- and WM-prompted trials. Search display condition was dummy-coded (Only = 0; Consistent = 1; Inconsistent = 2) so as to facilitate comparisons between Consistent and Only trials, as well as between Inconsistent and Only trials.

The maximal model (with all random effects specified) was singular, and thus we tested for over-parameterization with a principal component analysis (PCA) of the random-effects variance-covariance estimates using the `rePCA` function of the `lme4` package. This allowed us to determine the random effects that the data were capable of supporting (Bates et al., 2015). Excluded random-effects terms were removed on the basis of least explained variance until the model converged. In removing a random slope, it is assumed that the variable is invariant across participants. The Akeike Information Criterion indicated that the removal of terms identified by the PCA improved goodness of fit. Predictor estimate significance (based on conditional F-tests with Kenward-Roger approximation for the degrees of freedom) was computed using the `tab_model` function from the `sjPlot` package (Lüdtke, 2018; Wickham, 2018).

For visualization purposes, we show the RT difference between Consistent and Only trials, and between Inconsistent and Only trials; these difference scores were first calculated for each participant and then averaged across participants. We

also exponentiated model-predicted log RT to yield estimated RT effects in ms; these differences are reported for the conditions of interest.

Lastly, because generalized linear mixed models (GLMMs) can sometimes outperform linear mixed models on transformed data (Lo & Andrews, 2015), we re-ran all analyses from Experiments 1–4 with a generalized linear mixed model. This approach used the `glmer()` function from `lme4` in R (Bates et al., 2015; R Core Team, 2022). The RT data were fit with a gamma family distribution with a log link function. All of the GLMM analyses produced similar results to, and supported conclusions from, our main analysis on log RT. The gamma distribution, however, provided a numerically worse fit to our RT distributions than the normal distribution did for log RT; the log RT linear mixed model is therefore reported in the main text.

Results

Training and Testing Phases

Participants successfully learned the association between each scene and its color. In the Training Phase, mean (\pm SD) precision on the colorwheel (in degrees from the correct color; across all trials) generally improved over cycles: $6.72^\circ \pm 11.61^\circ$ for cycle one; $5.90^\circ \pm 11.78^\circ$ for cycle two; $5.59^\circ \pm 11.72^\circ$ for cycle three. The mean (\pm SD) number of cycles per scene needed to pass the Testing Phase was 1.45 ± 0.97 , and 75.39% of scene-color pairs only needed to be tested once. Mean (\pm SD) colorwheel precision was $3.57^\circ \pm 2.52^\circ$ from the correct hue on the final test of each scene (i.e., when reported correctly from LTM).

Search Phase

Search accuracy When searching for a color in LTM, participants successfully clicked the prompted color on the search display in $78.39\% \pm 11.19\%$ of trials. When searching for a color in WM, participants successfully clicked the prompted color in $89.15\% \pm 8.79\%$ of trials. No response was registered on the search display within the time allotted on $3.36\% \pm 2.53\%$ of LTM-prompted trials; on WM-prompted trials this rate was $1.36\% \pm 1.48\%$.

Unprompted color report Participants' unprompted color reports after the search display were considered accurate ($\leq 20^\circ$ from the correct hue) on $82.26\% \pm 10.69\%$ of trials when reporting a color from WM (i.e., on LTM-prompted trials). On trials in which participants were reporting the color in LTM (i.e., on WM-prompted trials), they were correct $88.65\% \pm 7.71\%$ of the time. When reporting the color from WM, $0.10\% \pm 0.32\%$ of trials did not receive a response in the time allotted; this never occurred when reporting the LTM color.

Search response times RTs and accuracy in the Search Phase are summarized in Table 1. Our main analysis investigated RTs in the search task as a function of prompt (WM vs. LTM) and search condition (Consistent, Inconsistent, Only). We only analyzed trials in which participants both (1) clicked on the correct color in the search display and (2) reported the unprompted color

accurately. The linear mixed model (see *Experiment 1/Data/Analysis*) revealed a main effect of prompt on log RTs such that participants were significantly slower to respond on the search display for LTM-prompted trials than for WM-prompted trials ($\beta = 0.080$, $SE = 0.007$, $p < 0.001$). See Fig. 3a for a plot of all fixed effect estimates, and Fig. 3b for a plot of raw RT by prompt.

Table 1 Summary of response times (RTs) and accuracy in the Search Phase

Prompted memory type	Search condition		Exp 1	Exp 2	Exp 3		Exp 4	
					Prompted shape	Prompted color	Prompted shape	Prompted color
Long-term memory	Consistent	Avg. RT (\pm SD)	1364.2 ms	1223.1 ms	1419.8 ms	1204.8 ms	1406.4 ms	1211.4 ms
		(Both-Correct)	(\pm 241.2)	(\pm 242.5)	(\pm 268.0)	(\pm 232.2)	(\pm 271.1)	(\pm 267.1)
		Avg. Search Accuracy (\pm SD)	79.3%	87.7%	82.9%	89.6%	83.3%	87.6%
			(\pm 14.4)	(\pm 10.0)	(\pm 16.1)	(\pm 11.3)	(\pm 15.8)	(\pm 13.8)
	Only	Avg. Unprompted Report Accuracy (\pm SD)	83.7%	82.3%	80.8%	78.1%	86.2%	84.6%
			(\pm 13.1)	(\pm 13.4)	(\pm 16.2)	(\pm 16.7)	(\pm 13.6)	(\pm 14.5)
		Avg. RT (\pm SD)	1410.2 ms	1261.2 ms	1469.6 ms	1252.4 ms	1428.1 ms	1219.6 ms
		(Both-Correct)	(\pm 249.1)	(\pm 256.9)	(\pm 262.4)	(\pm 280.5)	(\pm 290.8)	(\pm 295.7)
	Inconsistent	Avg. Search Accuracy (\pm SD)	77.9%	86.4%	83.3%	87.8%	80.9%	84.5%
			(\pm 13.3)	(\pm 11.4)	(\pm 13.5)	(\pm 11.5)	(\pm 16.5)	(\pm 15.0)
		Avg. Unprompted Report Accuracy (\pm SD)	80.1%	74.3%	74.7%	73.3%	82.2%	81.8%
			(\pm 13.8)	(\pm 15.4)	(\pm 17.7)	(\pm 18.7)	(\pm 16.0)	(\pm 16.6)
Working memory	Consistent	Avg. RT (\pm SD)	1416.9 ms	1257.0 ms	1508.6 ms	1259.1 ms	1470.2 ms	1182.7 ms
		(Both-Correct)	(\pm 244.7)	(\pm 246.3)	(\pm 254.6)	(\pm 277.9)	(\pm 294.6)	(\pm 284.2)
		Avg. Search Accuracy (\pm SD)	78.2%	86.9%	80.3%	87.1%	78.6%	84.4%
			(\pm 12.1)	(\pm 11.7)	(\pm 16.9)	(\pm 12.9)	(\pm 17.9)	(\pm 15.5)
	Only	Avg. Unprompted Report Accuracy (\pm SD)	82.6%	79.8%	77.3%	75.6%	81.1%	82.2%
			(\pm 12.2)	(\pm 13.3)	(\pm 16.9)	(\pm 16.6)	(\pm 16.2)	(\pm 16.0)
	Inconsistent	Avg. RT (\pm SD)	1149.9 ms	1022.4 ms	1225.0 ms	1121.5 ms	1180.7 ms	1018.6 ms
		(Both-Correct)	(\pm 164.9)	(\pm 177.8)	(\pm 196.8)	(\pm 221.8)	(\pm 211.4)	(\pm 204.7)
		Avg. Search Accuracy (\pm SD)	89.3%	94.5%	96.4%	96.4%	97.7%	98.0%
			(\pm 9.3)	(\pm 7.1)	(\pm 6.7)	(\pm 6.7)	(\pm 4.8)	(\pm 4.7)
	Only	Avg. Unprompted Report Accuracy (\pm SD)	88.5%	86.3%	90.5%	88.1%	89.4%	85.2%
			(\pm 9.9)	(\pm 11.7)	(\pm 10.8)	(\pm 12.6)	(\pm 11.2)	(\pm 13.1)
	Inconsistent	Avg. RT (\pm SD)	1184.6 ms	1051.3 ms	1296.5 ms	1161.4 ms	1220.8 ms	1032.2 ms
		(Both-Correct)	(\pm 189.3)	(\pm 192.8)	(\pm 208.1)	(\pm 269.0)	(\pm 226.1)	(\pm 225.9)
		Avg. Search Accuracy (\pm SD)	87.5%	95.4%	94.9%	94.4%	95.3%	96.7%
			(\pm 11.9)	(\pm 6.6)	(\pm 7.5)	(\pm 8.8)	(\pm 8.0)	(\pm 7.5)
	Inconsistent	Avg. Unprompted Report Accuracy (\pm SD)	88.4%	83.0%	85.9%	84.5%	80.2%	79.3%
			(\pm 9.8)	(\pm 13.4)	(\pm 14.2)	(\pm 16.7)	(\pm 15.8)	(\pm 16.7)
	Inconsistent	Avg. RT (\pm SD)	1183.6 ms	1043.6 ms	1328.9 ms	1140.0 ms	1298.2 ms	1016.5 ms
		(Both-Correct)	(\pm 172.0)	(\pm 184.1)	(\pm 223.9)	(\pm 230.2)	(\pm 231.5)	(\pm 206.2)
		Avg. Search Accuracy (\pm SD)	90.3%	93.7%	92.7%	93.6%	92.1%	94.5%
			(\pm 11.1)	(\pm 8.5)	(\pm 9.3)	(\pm 8.4)	(\pm 13.0)	(\pm 11.7)
	Inconsistent	Avg. Unprompted Report Accuracy (\pm SD)	88.8%	86.1%	87.4%	85.4%	83.5%	80.0%
			(\pm 10.5)	(\pm 11.0)	(\pm 13.2)	(\pm 13.9)	(\pm 15.2)	(\pm 16.3)

Mean RTs in the search task, accuracy in the search task, and accuracy in the unprompted feature report are shown separately for each prompted memory type, search condition, and prompted stimulus dimension (where relevant) in each experiment

As hypothesized, RTs were significantly faster on Consistent trials relative to Only trials ($\beta = -0.032$, $SE = 0.008$, $p < 0.001$; model-estimated RT benefit on Consistent vs. Only trials = 38.31 ms). Contrary to expectations, the model did not reveal an effect of Inconsistent (vs. Only) search display on log RT ($\beta = -0.004$, $SE = 0.008$, $p = 0.629$). There were no significant interactions between prompt and search condition. To illustrate these effects, the RT data are plotted in milliseconds by prompt and condition in Fig. 3c.

Discussion

Results from Experiment 1 indicated that visual search is facilitated when representations in LTM and WM could cooperate to support attention to a single location. Specifically, participants' RTs were faster when the LTM and WM colors were contained in the same circle in the search display, relative to when only one of the colors was present in the display. This attentional facilitation occurred regardless of whether the LTM or WM color was being searched for (i.e., was the prompted color), and, concomitantly, whether it was a WM or a LTM representation that was an accessory to the search task (i.e., was the unprompted color). However, when the LTM and WM colors were spatially distant from each other (i.e., were in different circles), and therefore would have competed – rather than cooperated – for attention, we found no evidence of attentional capture by the unprompted color, regardless of which memory type was prompted.

Our conclusion from Experiment 1 is that WM and LTM can both affect visual search in the same trial, but that their interaction was limited to when these memories could cooperate, and not found when they were in competition. This particular imbalance – cooperation but not competition – was not expected: We had hypothesized competition between

memories, potentially with WM competing with LTM-guided search more than the other way around. Accordingly, a second experiment was developed to confirm these findings and expand upon the phenomenon's generalizability. Three key points were addressed in this second experiment. First, to address concerns about high participant rejection and trial exclusion rates in Experiment 1, we reduced the demands placed on participants by shortening the length of each phase: The number of scene-color associations to be memorized (and later used in the Search Phase) on any given block was halved, and the number of blocks doubled to compensate. In addition, a feedback screen was added to the end of each search trial to encourage correct responses in the Search Phase. Second, a more perceptually uniform colorspace (CIE L*a*b*) was used in Experiment 2 to diminish any potential variance in discernibility between generated colors, enable the use of fewer constraints in selecting colors, and align more with prior literature. Lastly, the search display was reorganized to contain fewer elements, and these elements were moved closer to the center of the display to concentrate their presence in the foveal region of the visual field.

Experiment 2

Methods

Participants

Data were collected until the final sample comprised 115 participants who met the inclusion criteria. This sample size was selected to match that of the first experiment. To meet that target sample size, 148 participants were recruited for an online study using Prolific (www.prolific.co). As before, these

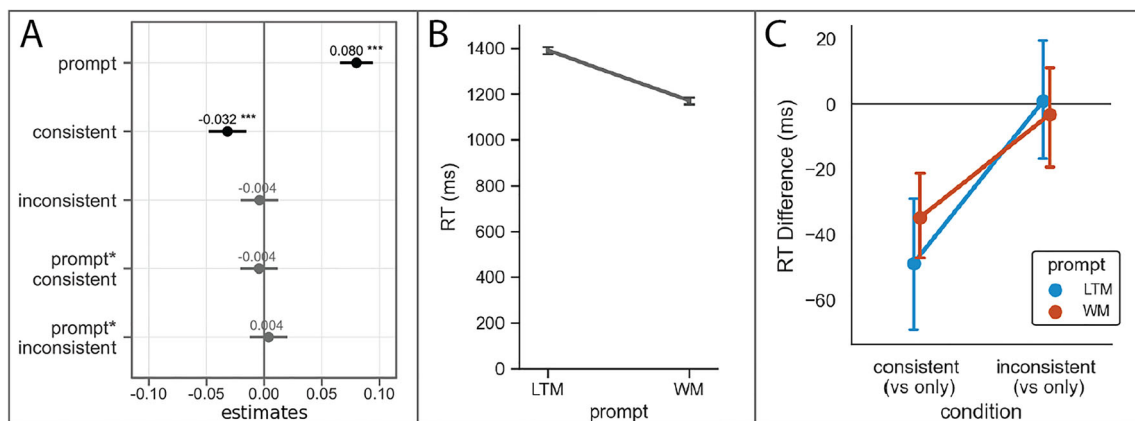


Fig. 3 Results from Experiment 1. **a** Coefficient estimates of the model for trial-wise log response time (RT). Error ribbons represent the 95% confidence interval; *** $p < .001$. **b** Long-term memory (LTM)-prompted trials (vs. working memory (WM)-prompted trials) were associated with a slower response on the search display. Error bars represent SEM of the within-participant (LTM-WM) difference. **c** Mean RT difference for

Consistent vs. Only and Inconsistent vs. Only trials. Search RTs were facilitated on Consistent (vs. Only) trials; RT on Inconsistent trials was no different from Only trials. Error bars represent SEM of the within-participant difference between Consistent and Only and Inconsistent and Only trials

participants were pre-screened for English fluency, nationality, and age; provided informed consent to a protocol approved by the Columbia University Institutional Review Board; and received \$6.50/h as compensation. Fourteen of these participants were unable to pass one of the Testing Phases and are not included in the final sample. Nineteen participants from the remaining sample did not correctly respond to $\geq 50\%$ of trials in the Search Phase (i.e., did not respond correctly to both the search display and unprompted color report), so none of their data are included. Our procedural changes were therefore successful at reducing the participant rejection rates seen in Experiment 1. Following these rejections, the final sample comprised 115 participants, as noted above ($M_{\text{age}} = 24.9 \pm 5.1$ years, $M_{\text{education}} = 15.3 \pm 2.5$ years). Ninety-three of these participants identified as women and four as non-binary; the remaining identified as men. Of this final sample, 83.3% identified as White, 14.0% as Asian, 6.1% as Black or African American, 1.8% as American Indian/Alaskan Native, 0.9% as Native Hawaiian or Pacific Islander, and 3.5% identified as part of a different racial group; in addition, 12.3% of these participants identified their ethnicity as Hispanic or Latino.

Stimuli

The methods for color assignment and scene stimuli were identical to those in Experiment 1.

Colorwheel The CIE $L^*a^*b^*$ colorwheel used in Experiment 2 was custom-built using JavaScript, but developed to function equivalently to the adapted HSL colorwheel from Experiment 1. The colorwheel was centered at $L^*=70$, $A^*=12$, $B^*=13$, with a radius of 60 (similar to prior online studies, e.g., Shin & Ma, 2016). Visual and textual feedback to responses on the colorwheel was identical to that in Experiment 1.

Color generation As in Experiment 1, two sets of five colors were generated, one to be assigned for LTM and one to be assigned for WM. This time, however, they were generated in CIE $L^*a^*b^*$ colorspace along the colorwheel described above (see Fig. 1b). CIE $L^*a^*b^*$ colorspace is roughly perceptually uniform and mimics the trichromatic vision of humans (i.e., is defined by luminance, green–red, and blue–yellow dimensions), meaning that two colors selected at a given distance apart are about as visually similar as two other colors selected at the same distance. Luminance was held constant so that stimuli sampled from the circumference of the colorwheel vary incrementally only in hue. Converting to this colorspace relaxed color generation constraints that were necessary when using the HSL colorspace of Experiment 1 (i.e., the use of the "center green" and its buffer zones; see *Experiment 1/Methods/Stimuli/Color generation*). Colors generated were therefore equidistant from one another on the CIE $L^*a^*b^*$ colorwheel, each color being

36° apart. Similar to Experiment 1, generated colors alternated between sets (i.e., no set contained colors within 76° of each other); a variable buffer (between 0° and 36°) was added at the "starting point" of the hue generation process to allow the generated sets to differ from one another across blocks; and assignment of sets to memory condition (WM or LTM) was randomized and balanced.

Procedure

As in Experiment 1, participants were directed via URL to Gorilla (www.gorilla.sc) and given instructions describing the structure of the experiment. These instructions differed minimally from those provided in Experiment 1 because, structurally, Experiment 2 was similar to Experiment 1 (see Fig. 4).

There were four key procedural differences between this and the first experiment. The first change is that blocks were split in half, relative to Experiment 1, such that there were six blocks comprising 15 trials rather than three blocks of 30 trials. This was done primarily to facilitate encoding of scene-color associations and reduce the failure rate in the Testing Phase, but also to potentially reduce errors from failures of sustained attention in the Search Phase. The two sets of colors were generated (see *Methods/Stimuli/Color generation*) at the beginning of every other block. Following the procedures in Experiment 1 (see *Experiment 1/Methods/Stimuli/Color assignment*), 30 trials were generated from those colors and then each 30-trial set was split randomly and equally between two blocks. In this way, the six blocks in Experiment 2 corresponded as closely as possible to the three in Experiment 1.

The second key change is that a re-study section was added between each cycle in the Testing Phase. This re-study section was equivalent to a mini-Training Phase for incorrectly responded-to scene-color pairs, and was implemented to enhance the encoding rate of scene-color pairs that participants were unable to recall on a given test cycle. An additional brief break of ≤ 33 s – like the break previously provided between each test cycle – was provided between re-study and test sections.

The third key change was to the search display (see Fig. 4c). The search display in Experiment 2 comprised three circles, yielding six differently colored half-circles. Concomitant with these modifications, fewer distractor colors were generated on each search trial. As such, distractors could be selected at a larger minimum colorspace distance from other elements on each trial's search display: Four or five distractor colors (depending on condition) were generated $\geq 50^\circ$ from the prompted/unprompted colors, and $\geq 35^\circ$ from one another. As in Experiment 1, the circle stimuli could not contain two colors $< 45^\circ$ apart, were at equal eccentricity from the center of the browser window, equidistant from one another, rotated around their central axes, and the circle stimuli were also

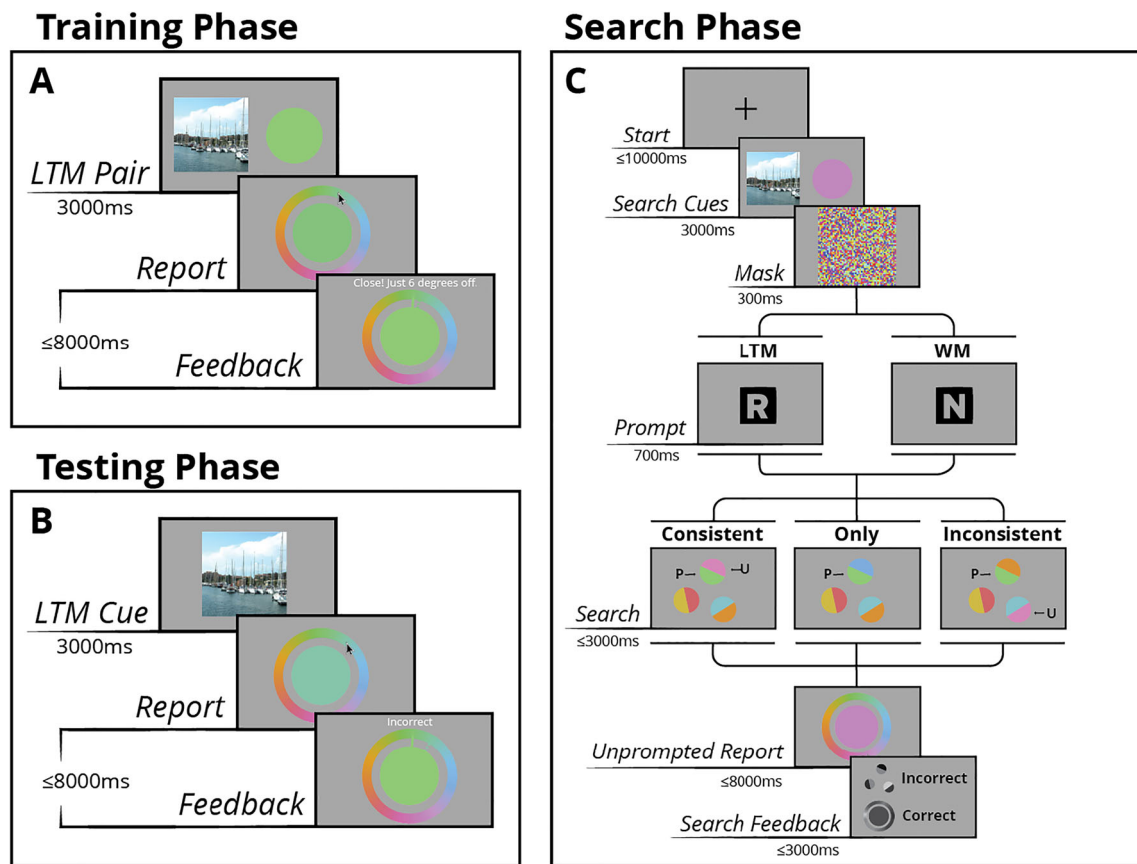


Fig. 4 Experiment 2 procedure is identical to Experiment 1 (see Fig. 2) except for the following: **a** 15 scene-color pairs were encoded in each block, rather than 30; there were six blocks rather than three; and the CIE $L^*a^*b^*$ colorspace was used. **b** A re-study section (not shown) helped participants re-encode scene-color pairs that had been imprecisely responded to. **c** Three circles were present in the search display instead of five, and the circles were closer to fixation. Lastly, participants were

provided with a Search Feedback screen, informing them of whether their responses on the search task and unprompted color report were correct / incorrect / not-responded-to. As before, the example search and unprompted color report displays in this figure are for a trial in which the long-term memory (LTM) ("remembered") color is prompted, and, for illustration purposes, the example search displays are labeled to denote the location of the prompted ("P") and unprompted ("U") colors

rotated around the center of the browser window in each search trial. The last change to the search display in Experiment 2 was that the radius of the (invisible) background circle around which the circle stimuli were placed was reduced by 55.56% to bring the color stimuli closer to the center of the browser window. This was done to ensure that particularities about the organization of the search display in Experiment 1 were not responsible for the imbalance in attentional guidance effects; and – because identification, discrimination, and popout of visual details, like color, are superior in the fovea than in the periphery (Aagten-Murphy & Bays, 2019; Gutwin et al., 2017) – it was supposed that moving the location of the prompted and unprompted colors closer to the center of the window might increase attentional capture effects.

The fourth and final change from Experiment 1 was that a feedback screen was added at the end of each search trial to inform participants of their performance (see the bottom of Fig. 4c for an illustration). Feedback for both the search task and the unprompted color report was "Correct," "Incorrect," or "No Response" (reports were considered incorrect if $\geq 20^\circ$ away on

the unprompted color report). This was done to encourage participants to respond correctly on both the search task and the unprompted color report, and thus to reduce the number of participant rejections and trial exclusions based on trial correctness.

Data

Trial exclusion criteria were identical to Experiment 1. Only those trials in which participants responded correctly to both the search display and to the unprompted color report were included. Across participants, an average of 77% of Search Phase trials were both-correct.

Analyses were identical to those conducted in Experiment 1.

Results

Training and Testing Phases

In the Training Phase, mean (\pm SD) precision on the colorwheel (in degrees from the correct color; across all trials)

generally improved over cycles: $8.53^\circ \pm 13.05^\circ$ for cycle one; $7.27^\circ \pm 9.82^\circ$ for cycle two; $7.17^\circ \pm 13.55^\circ$ for cycle three. The mean (\pm SD) number of cycles per scene needed to pass the Testing Phase was 1.63 ± 0.96 , and 60.18% of scene-color pairs only needed to be tested once. Mean (\pm SD) colorwheel precision was $4.20^\circ \pm 2.67^\circ$ from the correct hue on the final test of each scene (i.e., when reported correctly from LTM). Using the CIE L*a*b* colorspace — as opposed to HSL space used in Experiment 1 — therefore did not generally improve participants' precision on the colorwheel during the Training and Testing Phases. Nevertheless, participants successfully learned the association between each scene and its color.

Search Phase

Search accuracy When searching for a color in LTM, participants successfully clicked the prompted color on the search display in $87.00\% \pm 8.76\%$ of trials. When searching for a color in WM, participants successfully clicked the prompted color in $94.53\% \pm 5.63\%$ of trials. No response was registered on the search display within the time allotted on $2.13\% \pm 2.51\%$ of LTM-prompted trials; on WM-prompted trials this rate was $0.77\% \pm 1.06\%$.

Unprompted color report Participants' unprompted color reports after the search display were considered accurate ($\leq 20^\circ$ from the correct hue) on $78.80\% \pm 11.56\%$ of trials when reporting a color from WM (i.e., on LTM-prompted trials). On trials in which participants were reporting the color in LTM (i.e., on WM-prompted trials), they were correct $85.16\% \pm 9.23\%$ of the time. When reporting the color from WM, $0.24\% \pm 0.60$ of trials did not receive a response in the time allotted; this never occurred when reporting the LTM color.

Search response times RTs and accuracy in the Search Phase are summarized in Table 1. The primary analysis of RTs in the search task as a function of prompt (WM vs. LTM) and search condition (Consistent, Inconsistent, Only) revealed similar results as in Experiment 1. As before, we only analyzed “both-correct” trials in which participants both (1) clicked on the correct color in the search display and (2) reported the unprompted color accurately. The linear mixed model revealed a main effect of prompt on log RTs such that participants were significantly slower to respond on the search display for LTM-prompted trials than for WM-prompted trials ($\beta = 0.082$, $SE = 0.006$, $p < 0.001$). See Fig. 5a for a plot of all fixed effect estimates, and Fig. 5b for a plot of raw RT by prompt.

As in Experiment 1, RTs were significantly faster on Consistent trials relative to Only trials ($\beta = -0.027$, $SE = 0.008$, $p < 0.001$; model-estimated RT benefit on Consistent

vs. Only trials = 29.17 ms), and, again, there was no significant effect of Inconsistent (vs. Only) search display ($\beta = -0.008$, $SE = 0.008$, $p = 0.335$). Replicating Experiment 1, no significant interactions were revealed between prompt and search condition. To illustrate these effects, the RT data are plotted in milliseconds by prompt and condition in Fig. 5c.

Finally, we conducted an additional analysis to see if combining data across Experiments 1 and 2 would yield additional effects, given higher statistical power. To that end, we combined the data from Experiment 1 and Experiment 2 and then analyzed log RT in a similar multilevel linear mixed model (the only change being the addition of experiment version as an interacting fixed factor). We observed an effect of experiment version such that RTs were faster in Experiment 2 versus Experiment 1 ($\beta = -0.116$, $SE = 0.020$, $p < 0.001$). Significant effects of prompt (WM vs. LTM) and Consistent (vs. Only) search condition persisted in this combined model, but no other effect was yielded: No significant effect of Inconsistent (vs. Only) trial type, nor any significant interactions.

Discussion

Results from Experiment 2 replicated the findings from Experiment 1. Visual search continued to be speeded when the unprompted and prompted colors were in the same object, relative to when no unprompted color was present. As before, this attentional facilitation occurred regardless of whether the items being searched for were in LTM or WM, and, accordingly, whether the unprompted color was in WM or LTM, respectively. Likewise replicating Experiment 1, when the LTM and WM colors were organized on the search display such that they could have competed for attention, no evidence of attentional capture by the unprompted color was found, regardless of which memory type was prompted. The changes to block length and inclusion of a re-study section and feedback screen in Experiment 2 did, however, successfully ameliorate the high participant rejection and trial exclusion rates in Experiment 1.

Thus, the finding of cooperation, but not competition, between WM and LTM representations replicated in Experiment 2. This suggests that the particularities of the stimulus display or colorspace used in Experiment 1 were not responsible for these effects.

Participants were tasked with reporting the unprompted color at the end of each trial; this was done to encourage maintenance of the unprompted memory item over the course of the trial, including during the search task. It is possible, however, that individuals were able to temporarily suppress or reduce accessibility of the unprompted color during the search task and “reactivate” it during the unprompted report. If participants occasionally used the latter strategy, that may have reduced our likelihood of seeing competition. We therefore considered the possibility that unprompted memories

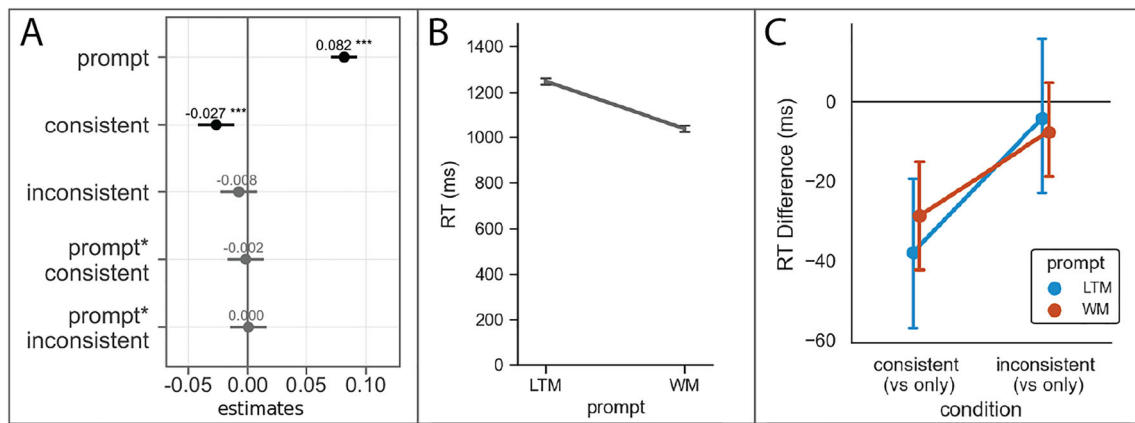


Fig. 5 Results from Experiment 2 were similar to those from Experiment 1. **a** Coefficient estimates of the model for trial-wise log response time (RT). Error ribbons represent the 95% confidence interval; *** $p < .001$. **b** Long-term memory (LTM)-prompted trials (vs. working memory (WM)-prompted trials) were associated with a slower response on the search display. Error bars represent SEM of the within-participant (LTM-WM)

may be more likely to capture attention if they are relatively actively represented and accessible. If this were the case, competition (Inconsistent (vs. Only) effect) may only have emerged in Experiments 1 and 2 under circumstances in which the unprompted memory was most readily accessible, i.e., when the unprompted report was executed relatively quickly. To test this, we first combined data across these Experiments to maximize power. Unprompted report RT was first z-scored within each relevant sub-condition (participant, prompt (WM or LTM), and search condition (Only, Inconsistent)) before being categorized by tertile (i.e., “fast”, “medium”, or “slow” unprompted report RT). The Inconsistent versus Only search RT difference was then computed separately for each participant and tertile. Finally, we tested to see if competition (the Inconsistent (vs. Only) effect) significantly deviated from zero when participants’ unprompted report RT was in the fastest tertile with a one-sample t-test; however, this effect was not statistically significant ($t(229) = 1.80$, $p = 0.41$). This indicates that the unprompted item did not capture attention in Experiments 1 and 2, even when participants’ unprompted report responses were quickest.

Why did we fail to find competition between WM and LTM during attentional guidance? One possibility – inspired by research on redundancy gains and coactivation models (Miller, 1982; Mordkoff & Yantis, 1993) – is the presence of an object-wise winner-takes-all cognitive operation (Koch & Ullman, 1985). That is, perceptual evidence may accumulate from the location of both the prompted and the unprompted colors in the search display, but this evidence may only be summed (or otherwise combined) when it accumulates from a single object (here, a circle; Danek & Mordkoff, 2011; van Ede & Nobre, 2022; see also Treisman & Gelade, 1980). This would yield faster RTs on Consistent vs. Only trials because evidence is combined across the prompted and unprompted colors in the

difference. **c** Mean RT difference for Consistent vs. Only and Inconsistent vs. Only trials. Search RTs were facilitated on Consistent (vs. Only) trials; RT on Inconsistent trials was no different from Only trials. Error bars represent SEM of the within-participant difference between Consistent and Only and Inconsistent and Only trials

same circle, resulting in a decision criterion being reached sooner. But this process would yield no difference in RTs between Inconsistent and Only trials, because the evidence accumulated for the unprompted color, hypothesized to be weaker than that for the prompted color, arrives from a different object and hence may lose the winner-takes-all competition.

To explore whether this potential winner-takes-all mechanism is a general phenomenon, we conducted a third experiment in which the WM and LTM representations came from different stimulus dimensions: A color and a shape rather than two colors. In addition to allowing us to test the generalizability of this proposed mechanism, using color and shape enabled us to explore more unified objects, in which the two dimensions resided in the same physical space (i.e., “square” and “red” would become a red square, as opposed to two colors in two distinct halves of a circle).

To this end, we leveraged recent work by Li et al. (2020), who created a validated circular shape (VCS) space (see *Experiment 3/Methods/Stimuli/Colorwheel and shapewheel*). Using the VCS space allowed us to precisely manipulate the perceived visual similarity of shape, analogously to how we used CIE L*a*b* colors and their associated colorwheel. This functionally enables the use of a two-dimensional perceptually uniform space in Experiment 3.

Experiment 3

Methods

Participants

Data were collected until the final sample comprised 192 participants who met the inclusion criteria. To meet that target sample

size, 253 participants were recruited for an online study using Prolific (www.prolific.co). As before, these participants were pre-screened for English fluency, nationality, and age; provided informed consent to a protocol approved by the Columbia University Institutional Review Board; and received \$6.50/h as compensation. Using the same rejection criteria as in Experiments 1 and 2: 26 of these participants were unable to pass one of the Testing Phases and 35 participants of the remaining sample did not correctly respond to $\geq 50\%$ of the search trials' search displays and unprompted reports, so none of their data are included. Following these rejections, the final sample comprised 192 participants, as noted above ($M_{\text{age}} = 26.3 \pm 5.7$ years, $M_{\text{education}} = 15.2 \pm 2.1$ years).

This sample size was selected to match Experiments 1 and 2 with regards to the number of observations per condition, across participants (power being a joint function of sample size and number of trials; Baker et al., 2021). There were double the number of conditions in Experiment 3 relative to Experiments 1 and 2, to accommodate the addition of a new stimulus dimension. Because an additional 18 trials were added per participant in Experiment 3 (for counterbalancing purposes), we did not quite need to double the sample size: $(115 \text{ participants} * 90 \text{ trials})/\text{six conditions}$ (in Experiments 1 and 2) $= (192 \text{ participants} * 108 \text{ trials})/\text{12 conditions}$ in Experiment 3.

One hundred and thirty-seven of the participants in the final sample identified as women, six as non-binary or otherwise gender non-conforming, one did not report a gender, and the remaining identified as men. Of this final sample, 82.0% identified as White, 11.6% as Asian, 7.4% as Black or African American, 2.1% as American Indian/Alaskan Native, 0.5% as Native Hawaiian or Pacific Islander, and 3.2% identified as part of a different racial group; in addition, 12.3% of these participants identified their ethnicity as Hispanic or Latino.

Stimuli

Colorwheel and shapewheel The same CIE $L^*a^*b^*$ colorwheel from the prior experiment was used in Experiment 3 (see *Experiment 2/Methods/Stimuli/Colorwheel*).

To enable the use of a shape dimension, we used the Validated Circular Shape (VCS) space (Li et al., 2020), which is comparable to the previously-used CIE $L^*a^*b^*$ colorspace: Angular distance along a circle functions as a proxy for perceived visual similarity, but for shape instead of color. The VCS shapewheel was custom-built using JavaScript to function analogously to the CIE $L^*a^*b^*$ colorwheel. First, the 360 VCS shape images were batch-processed using a custom preset with the Image Trace tool in Adobe Illustrator to convert them to scalable vector graphics (SVG) files, an extensible markup language- (XML-) based vector format optimized for display in the browser – thus, the shapes can be rendered at any size without loss of quality, like the colors and circles in

Experiments 1 and 2. Each shape's underlying XML was then inserted into HTML `<svg>` containers.

The shapewheel itself was a black circle with no indicators as to where individual shapes were located. Like the colorwheel, the shapewheel contained 360 segmentations and was rotated randomly on each trial ($0\text{--}359^\circ$) around its central axis so that location on the shapewheel could not function as a proxy for shape. Similar to the colorwheel, participants could hover their cursor over the shapewheel to view the shape stimulus, which was presented in black in the center of the circle. That is, participants moved their cursor around the shapewheel to find where the shape they wished to report was located on any given trial.

Upon clicking a point on the shapewheel, the corresponding shape was selected and five visual feedback elements were placed on or in the shapewheel: (1) The shape displayed centrally was replaced with the correct shape for that trial; (2) the correct shape was also displayed in gray inside of the black outline of the shapewheel circle, at the position on the shapewheel that the correct shape was located; (3) a solid line, like in the colorwheel, was drawn in black from the center of the circle to the correct position; (4) a dotted line, as for the colorwheel, was drawn from the center of the circle to the corresponding selected shape on the shapewheel; and (5) that selected shape was displayed in gray inside of the black outline of the shapewheel circle at the location clicked, so that participants could see the exact shape they had selected. This last piece of visual feedback (5) was only displayed, however, when a participants' guess was $\geq 10^\circ$ away from the correct shape, to avoid overlap of the gray shape elements, which would have rendered them uninformative (see Fig. 6 for an example of these visual feedback elements).

Lastly, like with the colorwheels, textual feedback was provided regarding precision on the shapewheel. This textual feedback was identical to that from the colorwheels in Experiments 1 and 2 (see *Experiment 1/Methods/Procedure*).

All of these feedback elements were seen on the Training and Testing Phase report shapewheels. On the shapewheel used in the Search Phase, however – on the unprompted report screen – only visual feedback element 4 was used. Thus, on the Search Phase shapewheel no indication was given as to the correct shape when a shape was selected; feedback about shapewheel precision (accurate or inaccurate) was provided only on the search feedback screen directly after the unprompted report, to match Experiment 2.

Color and shape generation As in Experiment 2, two sets of colors were generated on the same CIE $L^*a^*b^*$ colorwheel, one to be assigned for LTM and one to be assigned for WM. This time, however, only six colors were generated total, at equidistant intervals (60°) from one another (see Fig. 1c, top). As before: Colors alternated being assigned to one set or the

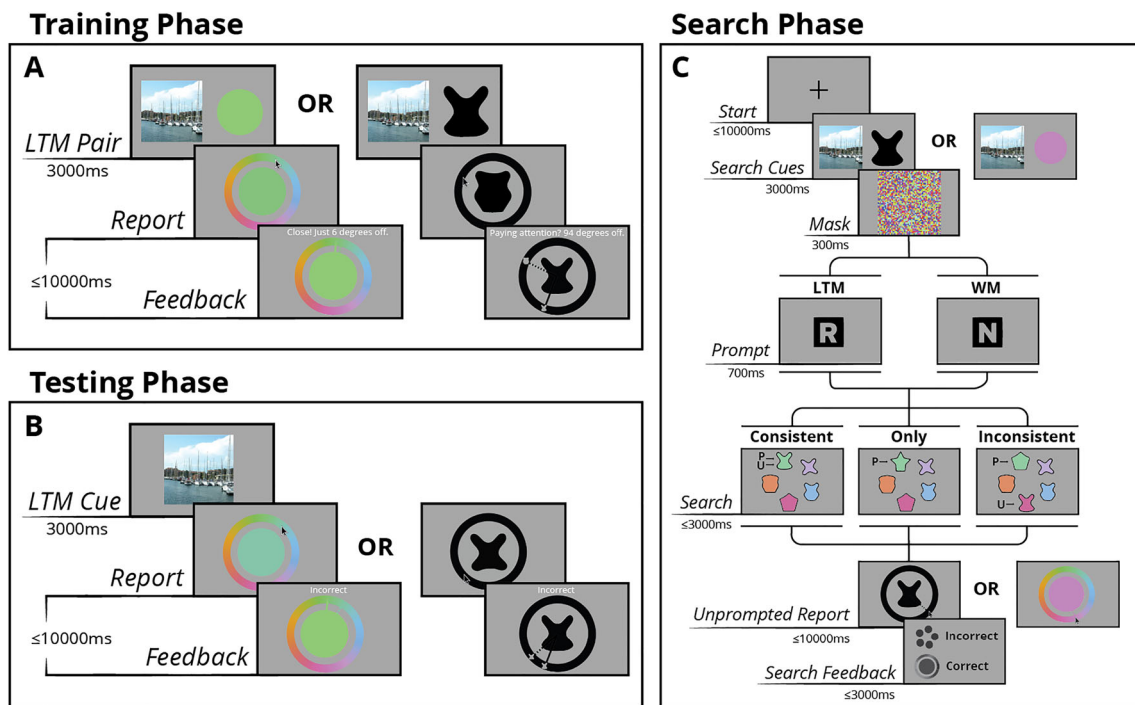


Fig. 6 Experiment 3 procedure is identical to Experiment 2 (see Fig. 4) except for the following. **a** Scene-stimulus pairs were encoded as before, but each scene could either be paired with a color (left) or a shape (right). There were now 18 scene-stimulus pairs per block (across six blocks), and participants were given ≤ 10 s on the report screen. **b** All 18 scene-stimulus pairs were tested per block, and participants were given ≤ 10 s on the report screen. **c** Search cues comprised one long-term memory (LTM) stimulus and one working memory (WM) stimulus, as before, but each scene that had previously been paired with a color was now paired with a shape WM cue (left); and each scene that had previously been paired with a shape was now paired with a color WM cue (right). Five shapes were presented on the search display, each filled with a solid color and outlined with black. Similar to previous experiments, there were three conditions for this search display: *Consistent* trials contained both the prompted and unprompted stimuli in a single, unitized object (i.e., the prompted color in the unprompted shape, or vice versa); *Only* trials contained the prompted stimulus only (and not the unprompted stimulus); and *Inconsistent* trials

contained both the prompted and unprompted stimuli, but these two stimulus features were not unitized, instead, they were separated and in different locations on the display. After the search display, participants were asked to report the unprompted shape on the shapewheel (left) or the unprompted color on the colorwheel (right). As before, the example search and unprompted report displays in this figure illustrate a trial in which the LTM ("remembered") color is prompted. For illustration purposes, the example search displays are again labeled to denote the location of the prompted ("P") and unprompted ("U") stimuli. This example trial is also a trial in which a color (retrieved from LTM) was the prompted stimulus; the use of "OR" in this figure represents what each screen would look like if this example trial had a prompted shape (retrieved from LTM), for each screen that would have differed, except search displays. Note that the example scene image is shown "paired" with a color and shown "paired" with a shape strictly for illustration purposes; no scene was actually paired with more than one stimulus feature

other (i.e., no set contained colors within 120° of each other); a variable buffer (between 0° and 60°) was added at the "starting point" of the hue generation process to allow the generated sets to differ from one another across blocks; and assignment of color sets to memory condition (WM or LTM) was randomized and balanced.

In addition, two sets of shapes were generated on the VCS shapewheel, one to be assigned for LTM and one to be assigned for WM, in the same way that colors were generated for this Experiment (see above; Fig. 1c, bottom). That is: Six shapes equidistant from one another (60° apart) alternated in assignment to one set or the other (i.e., no set contained shapes within 120° of each other); the same (0 – 60°) buffer was added to the "starting point" of the shape generation process; and assignment of shape sets to memory condition (WM or LTM) was randomized and balanced.

Color and shape assignment Two stimuli were still cued on each search trial, one color and one shape: One LTM color or shape (that would be retrieved from LTM) and one WM shape or color (that would be presented on the screen). Beyond having one LTM item and one WM item, there were no other constraints placed on which colors/shapes could be combined on a given search trial, permitting the use of 18 possible color/shape pairs.

To match Experiment 2 as closely as possible, blocks continued to be split in half such that new shape and color stimuli were generated (see *Methods/Stimuli/Color and shape generation*) every other block (i.e., each 36-trial set was split in half between two blocks). Shape/color pairs were then assigned to satisfy four balancing conditions: (1) Each half-block had an equal number of LTM- and WM-prompted trials, (2) there was an equal number of color- and shape-prompted

trials per half-block, and (3) there was an equal number of Consistent, Inconsistent, and Only trials in each half-block. Once these parameters had been established for color and shape assignment, it was algorithmically impossible to ensure that each shape and color was displayed exactly an equal number of times (three) per half-block. Consequently, (4) each shape and each color was required to be present in each half-block at least twice (i.e., a given color or shape was sometimes displayed only two times in a given half-block and sometimes displayed four times). Therefore, across every whole block (of 36 trials), each color/shape pair was cued twice, and each color and each shape was cued six times.

Scenes The same scenes were used in Experiment 3 as in Experiments 1 and 2. An additional 18 scenes were selected from the same scene stimulus set – in the same way as before (i.e., avoiding substantial visual overlap and text) – to account for the increase in number of trials, and randomized as before (see *Experiment 1/Methods/Stimuli/Scenes*).

Procedure

As before, participants were first directed via URL to Gorilla (www.gorilla.sc) and given instructions describing the structure of the experiment. These instructions differed minimally from those provided in Experiments 1 and 2 because, structurally, the procedure of Experiment 3 was designed to be as close as possible to that used in Experiment 2 (see Fig. 6). The sole change, albeit a substantial one, was the inclusion of a second dimension for all the memory and search stimuli (i.e., shape, in addition to color). This change had an effect on each phase of the experiment, but the structure of these phases remained unchanged from Experiment 2.

In each Training Phase, there were 18 scene-stimulus pairs, half of which were scene-color pairs and half of which were scene-shape pairs. The time allotted for responding on the colorwheel or shapewheel on the response screen was increased by two seconds relative to Experiment 2 (to ≤ 10 s; see Fig. 6a). This was done to allow individuals the additional time needed to select a shape on the shapewheel, because there was no visual cue on the shapewheel as to the location of a given shape until participants used their cursor to explore where the shapes were (see Fig. 6a). No other changes were made to the Training Phase. The Testing Phase was also similar to Experiment 2. Participants were tested on the 18 scene-stimulus pairs from the prior Training Phase, and given the same amount of time (≤ 10 s) on the report screen as in the Training Phase (see Fig. 6b).

The Search Phase in Experiment 3 was modified to accommodate the added stimulus dimension as well (see Fig. 6c). As described above (*Methods/Stimuli/Color and shape assignment*), each of the 18 scenes that had been associated

with a shape in the Testing/Training Phase was now paired with a WM color cue, and vice versa, in the search cues screen. As in Experiments 1 and 2, each scene was shown once. Upon being prompted with the same “R” (“remembered”; LTM) or “N” (“new”; WM) stimuli as in Experiments 1 and 2, participants were presented with a search display that was modified to integrate the additional shape dimension: The search display now comprised five shapes, each outlined with black and filled with a solid color. On Consistent search displays the prompted color or shape was unitized with the unprompted shape or color, such that the prompted and unprompted stimulus features were bound to a single object. On Only search displays the prompted stimulus feature (either color or shape) was present on the display, but the unprompted stimulus feature was absent entirely. On Inconsistent search displays the prompted and unprompted stimulus features were both present, but in separate locations on the screen; they were not unitized into one object. Instead, both the prompted and unprompted stimulus features were unitized with a distractor feature. These distractor features were generated at random, but at a minimum distance of 60° from both the LTM and WM stimulus features cued on that trial, and at a minimum distance of 45° from other distractors. After clicking the prompted stimulus feature, participants were instructed to report the unprompted stimulus feature on either the colorwheel or the shapewheel before being provided feedback about their accuracy, as in Experiment 2.

Data

Trial exclusion criteria were identical to Experiments 1 and 2. Only those trials in which participants responded correctly to both the search display and to the unprompted color report were included. Across participants, an average of 76% of Search Phase trials were both-correct.

Analyses were similar to those conducted in Experiments 1 and 2. We used the same analysis tools, and, as before, modeled the primary dependent variable, log-transformed RT on the search task, with multilevel linear regression on participants’ trial-wise data. We fit an analogous mixed-effects model, but included one more interaction term – stimulus dimension (shape or color, effect-coded as 1 and -1 respectively) – as a fixed factor; we also included this as a random effect term to allow the estimates to vary between participants. That is: $\text{lmer}(\log(\text{RT}) \sim \text{prompt} * \text{searchCondition} * \text{stimulusDimension} + (1 + \text{prompt} + \text{stimulusDimension} | \text{Participant}))$.

Results

Training and Testing Phases

Participants successfully learned the association between each scene and its stimulus, and precision when reporting shapes on

the shapewheel was fairly comparable to reporting colors on the CIE $L^*a^*b^*$ colorwheel. In the Training Phase, mean (\pm SD) precision on the color/shapewheel (in degrees from the correct color/shape) generally improved over cycles: $8.16^\circ \pm 11.31^\circ$ for colors and $9.62^\circ \pm 15.30^\circ$ for shapes in cycle one; $6.93^\circ \pm 10.06^\circ$ for colors and $7.24^\circ \pm 11.25^\circ$ for shapes in cycle two; $6.40^\circ \pm 8.29^\circ$ for colors and $6.78^\circ \pm 10.76^\circ$ for shapes in cycle three.

The mean (\pm SD) number of cycles per scene needed to pass the Testing Phase was 1.45 ± 0.81 (for color-associated scenes, 1.49 ± 0.85 ; for shape-associated scenes, 1.42 ± 0.78), and 68.56% of scene-stimulus pairs only needed to be tested once (for color-associated scenes, 67.21%; for shape-associated scenes, 69.91%). Mean (\pm SD) precision on the color/shapewheel was $3.95^\circ \pm 2.62^\circ$ from the correct stimulus on the final test of each scene (i.e., when reported correctly from LTM): $4.06^\circ \pm 2.66^\circ$ for colors and $3.83^\circ \pm 2.58^\circ$ for shapes.

Search Phase

Search accuracy When searching for a stimulus feature in LTM, participants successfully clicked the prompted feature on the search display in $85.19\% \pm 8.58\%$ of trials, and when searching for a feature in WM, the rate was $94.72\% \pm 4.47\%$ of trials. When searching for a color, participants successfully clicked the prompted feature on the search display in $91.48\% \pm 5.82\%$ of trials, and when searching for a shape, the rate was $88.43\% \pm 6.89\%$ of trials. No response was registered on the search display within the time allotted on $3.79\% \pm 2.81\%$ of LTM-prompted trials; on WM-prompted trials this rate was $1.62\% \pm 1.63\%$. On trials in which the prompted stimulus was a color, this rate was $2.32\% \pm 1.92\%$; and on trials with prompted shapes, this rate was $3.09\% \pm 2.39\%$.

Unprompted (color or shape) report Participants' unprompted reports after the search display were considered accurate ($\leq 20^\circ$ from the correct color or shape) on $76.64\% \pm 11.71\%$ of trials when reporting from WM (i.e., on LTM-prompted trials). On trials in which participants were reporting from LTM (i.e., on WM-prompted trials), they were correct $86.97\% \pm 8.72\%$ of the time. When reporting a shape (i.e., on trials with a prompted color), $80.84\% \pm 10.98\%$ of trials were correct. When reporting a color (i.e., on trials with a prompted shape), this rate was $82.76\% \pm 9.71\%$. When reporting the stimulus from WM, $0.18\% \pm 0.54\%$ of trials did not receive a response in the time allotted; this never occurred when reporting from LTM. When reporting a color, this rate was $0.24\% \pm 0.59\%$; and when reporting a shape, it was $0.15\% \pm 0.46\%$.

Search response times RTs and accuracy in the Search Phase are summarized in Table 1. As before, we only analyzed trials in which participants both (1) clicked on the correct stimulus

feature in the search display and (2) reported the unprompted stimulus feature accurately. The primary analysis of RTs in the search task — as a function of prompt (WM vs. LTM), search condition (Consistent, Inconsistent, Only), and stimulus dimension (Shape, Color) — revealed, as before, a main effect of prompt on log RTs such that participants were significantly slower to respond on the search display for LTM-prompted trials than for WM-prompted trials ($\beta = 0.048$, $SE = 0.005$, $p < 0.001$). See Fig. 7a for a plot of all fixed-effect estimates, and Fig. 7b for a plot of raw RT by prompt. In addition, a main effect of stimulus dimension was found, indicating that participants were faster to respond on trials in which the prompted stimulus was a color rather than a shape ($\beta = -0.079$, $SE = 0.005$, $p < 0.001$).

As before, RTs were significantly faster on Consistent trials relative to Only trials ($\beta = -0.036$, $SE = 0.006$, $p < 0.001$; model-estimated RT benefit on Consistent vs. Only trials = 42.98 ms). This benefit on Consistent trials, however, interacted with stimulus dimension: When the prompted stimulus feature was a shape (vs. a color), there was a larger effect of a trial being Consistent (relative to Only; $\beta = 0.013$, $SE = 0.006$, $p = 0.023$). That is, the effects of having a Consistent (vs. Only) search display were greater when color was the accessory dimension (and shape the target) rather than when shape was the accessory dimension (and color the target).

Unlike in Experiments 1 and 2, however, there was a significant effect of Inconsistent (vs. Only) search display ($\beta = 0.013$, $SE = 0.006$, $p = 0.031$; model-estimated RT cost = 15.80 ms). This Inconsistent predictor interacted with stimulus dimension such that the effects of having an Inconsistent (vs. Only) search display were significantly larger when color (rather than shape) was the accessory dimension ($\beta = -0.013$, $SE = 0.006$, $p = 0.028$). Indeed, competition (on Inconsistent vs. Only trials) was reliable only when color was the accessory item ($\beta = -0.026$, $SE = 0.008$, $p = 0.001$) and not when shape was the accessory item ($\beta = -0.001$, $SE = 0.009$, $p = 0.952$). Thus, both cooperation (on Consistent vs. Only trials) and competition (on Inconsistent vs. Only trials) were greater when color was the accessory dimension, rather than shape.

There was one final interaction between stimulus dimension and prompt. The RT benefit conferred by searching for a WM (vs. LTM) stimulus feature was larger when that stimulus feature was in the shape (vs. color) dimension ($\beta = -0.010$, $SE = 0.006$, $p = 0.004$). As in the first two experiments, no interactions were revealed between prompt and search condition, nor were there any three-way interactions. To illustrate these effects, the RT data are plotted in milliseconds by prompt and condition in Fig. 7c, and by stimulus dimension and condition in Fig. 7d.

Discussion

Our conclusion from Experiment 3 is threefold. First, WM and LTM representations continued to cooperate in visual

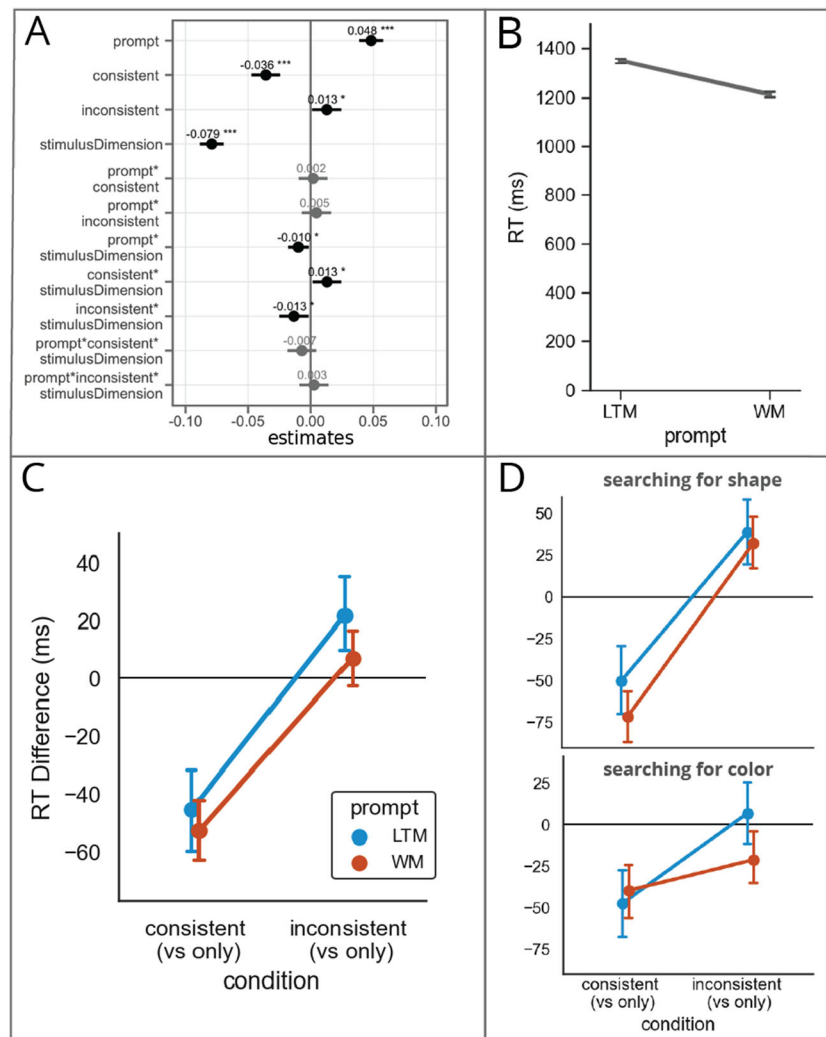


Fig. 7 Results from Experiment 3. **a** Coefficient estimates of the model for trial-wise log response time (RT). Error ribbons represent the 95% confidence interval; *** $p < .001$, * $p < .05$. **b** Long-term memory (LTM)-prompted trials (vs. working memory (WM)-prompted trials) were associated with a slower response on the search display. Error bars represent SEM of the within-participant (LTM-WM) difference. **c** Mean RT difference by prompt for Consistent (vs. Only) and Inconsistent (vs. Only) trials. Search RTs were facilitated on Consistent (vs. Only) trials, and delayed on Inconsistent (vs. Only) trials. Error bars represent SEM of the

within-participant difference between Consistent and Only and Inconsistent and Only trials. **d** Mean RT difference by dimension of the prompted stimulus for Consistent (vs. Only) and Inconsistent (vs. Only) trials. The model output – represented in (A) – confirms the interaction visualized here between stimulus dimension and search condition: The effects of a trial being Consistent (vs. Only) or Inconsistent (vs. Only) were larger for trials in which participants searched for a shape (and thus the unprompted stimulus was a color). Error bars represent SEM of the within-participant difference between Consistent and Only and Inconsistent and Only trials

search, producing behavioral facilitation when the activated features were in the same object. This occurred both when the unprompted feature was a shape and when it was a color. Second, in contrast to Experiments 1 and 2, we found that, under certain task circumstances, WM and LTM can compete, with the unprompted memory capturing attention even when it hinders performance; this effect only occurred, however, when the unprompted feature was a color (and the prompted feature a shape), and not vice versa.

That we observed competitive effects only when a color was the unprompted feature is curious, as no competitive effect was yielded in Experiments 1 and 2 despite the fact that, in those experiments, color was the sole stimulus dimension, and therefore was always the dimension of the unprompted feature. We will explore possible explanations for this in the *General discussion*.

The third conclusion from Experiment 3 is that the source of a memory during memory-guided attention (whether WM or LTM) did not interact with these cooperative or competitive

effects. This finding contradicts a potential asymmetry in memory-guided attention, mentioned in the Introduction, in which accessory (unprompted) WMs may compete with LTM-guided search more than the other way around.

Why did we not see differential cooperation or competition between WM and LTM? One possibility is that, during this task, unprompted LTM stimuli were retrieved and subsequently represented in the direct access region of WM – just as an unprompted WM representation may be – rather than being represented in a distinct store, like activated LTM. That is, perhaps we see no difference between LTM and WM because both representations are in the same cognitive store and thus operate on attention via the same cognitive mechanism. If it is indeed the case that the prompted memory is always in the focus of attention and the unprompted memory is always in the direct access region of WM, then cooperation and competition between two WM representations should not be different from cooperation and competition between two LTM representations. A fourth experiment was conducted to test this idea. Experiment 4 was designed to be as similar to Experiment 3 as possible, but instead of pairing each LTM (shape or color) with a WM (color or shape), we instead paired LTM stimuli together and WM stimuli together (i.e., each LTM color was paired with a LTM shape, and each WM color was paired with a WM shape). If unprompted LTM features are indeed held in the direct access region, LTM trials and WM trials in Experiment 4 should appear identical.

Experiment 4

Methods

Participants

Data were collected until the final sample comprised 192 participants who met the inclusion criteria, to match the sample size of Experiment 3. To meet that target sample size, 274 participants were recruited for an online study using Prolific. As before, participants were pre-screened for English fluency, nationality, and age; provided informed consent to a protocol approved by the Columbia University Institutional Review Board; and received \$6.50/h as compensation. Thirty-two of these participants were unable to pass one of the Testing Phases and 50 participants from the remaining sample did not get $\geq 50\%$ of search trials both-correct, so none of their data are included.

Following these rejections, the final sample comprised 192 participants, as noted above ($M_{\text{age}} = 27.7 \pm 6.2$ years, $M_{\text{education}} = 15.3 \pm 2.4$ years). 120 of these participants

identified as women, 16 as non-binary or otherwise gender non-conforming, two did not report a gender, and the remaining identified as men. Of this final sample, 89.8% identified as White, 7.0% as Asian, 4.3% as Black or African American, 2.1% as American Indian/Alaskan Native, 0% as Native Hawaiian or Pacific Islander, and 2.1% identified as part of a different racial group; in addition, 14.4% of these participants identified their ethnicity as Hispanic or Latino.

Stimuli

All stimuli and stimulus creation procedures were as described in Experiment 3 (see *Experiment 3/Methods/Stimuli*; for color and shape generation, see Fig. 1d).

Procedure

As before, participants were directed via URL to Gorilla (www.gorilla.sc) and given instructions describing the structure of the experiment (see Fig. 8). The Training Phase (Fig. 8a) and Testing Phase (Fig. 8b) remained unchanged from Experiment 3: Participants learned and then completed tests on associations between scene images and stimuli (colors or shapes) until they could report each stimulus correctly.

The Search Phase in Experiment 4 was modified from the previous experiment to accommodate the pairing of LTM stimuli to other LTM stimuli, and the pairing of WM stimuli to other WM stimuli (see Fig. 8c). As in Experiments 1–3, participants first clicked a fixation cross to begin each trial before being shown two search cues. In Experiment 4, these search cues were either two LTM cue scenes, one of which had been associated with a color, the other with a shape (LTM trials); or two WM cues, one color and one shape, presented directly on the screen (WM trials). The “R” and “N” stimuli from Experiments 1–3 that previously indicated whether the prompted item on any given trial was the “remembered” stimulus (LTM-prompted trials) or the “new” stimulus (WM-prompted trials) could no longer be used, as each trial in Experiment 4 had either two “remembered” or two “new” stimuli. Thus, participants were now prompted to either search for the stimulus associated with the left or the right side of the search cues screen. This prompt took the form of a square black outline filled in with black on the left side or the right side, respectively. LTM trials were now those in which participants searched for the stimulus feature associated with the scene presented on the left or the right side of the screen; and WM trials were those in which participants searched for the stimulus feature presented directly on the left or the right side of the screen. Upon being prompted with the left or right prompt, participants were presented with a search display identical to the one used in Experiment 3. As before, after

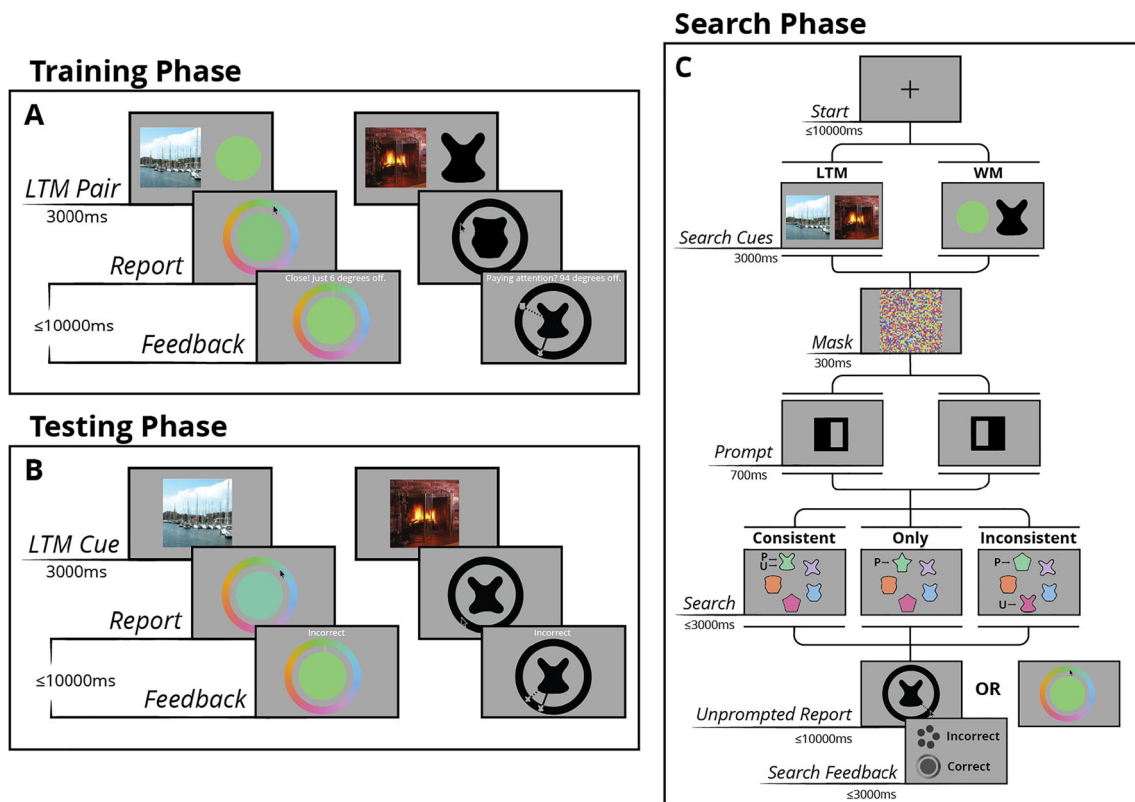


Fig. 8 Experiment 4 procedure. Training Phase (a) and Testing Phase (b) procedures were identical to those in Experiment 3 (see Fig. 6). The Search Phase (c) was identical to Experiment 3, except for the following changes. Search cues now comprised either two scene images, one of which had been previously associated with a color and the other a shape (long-term memory (LTM) trials); or one color and one shape presented directly on the screen (working memory (WM) trials). The prompt now indicated whether a given trial's target stimulus feature was associated with the left or the right side of the search cues screen. In this example trial, the left prompt would indicate a search for green, regardless of whether it was a LTM or WM trial (but note that LTM colors and shapes were different from WM colors and shapes in each block of the actual

experiment; the colors and shapes are the same here for visualization purposes). The search and unprompted report screens were identical to that of Experiment 3, as was the search feedback screen. The example search displays in this figure illustrate an LTM trial in which the left scene is prompted. For illustration purposes, the example search displays are labeled to denote the location of the prompted ("P") and unprompted ("U") features. This example trial is also a trial in which a color (retrieved from LTM) was the prompted stimulus; the use of "OR" in this figure represents what the unprompted report screen would look like if this example trial had a prompted shape (i.e., if the right prompt was presented)

clicking the prompted stimulus feature, participants were instructed to report the unprompted stimulus feature on either the colorwheel or the shapewheel (i.e., if prompted to search for a color, participants reported the unprompted shape, and vice versa), before being provided feedback about their accuracy.

Data

Trial exclusion criteria were identical to Experiments 1–3: Only those trials in which participants responded correctly to both the search display and to the unprompted color report were included. Across participants, an average of 76% of Search Phase trials were both-correct.

We initially modeled the data with the addition, relative to Experiment 3, of one variable. Unlike the previous experiments, the search prompt was directional (i.e., prompted the

participant to search for features associated with the left or right side of the search cues screen; Fig. 8c). We therefore added a "target-side agreement" variable (effect-coded: 1 = same and -1 = different) to control for differences in RT based on whether the prompted feature in the search display appeared on the same versus different side as the side indicated by the prompt. We had expected that participants may respond faster when there is agreement between the location of the search target (i.e., whether on the left or right side) and the side of the screen prompted. Participants indeed responded faster when the search target was on the same versus different side of the screen as that indicated by the prompt ($\beta = -0.048$, $SE = 0.008$, $p < 0.001$). However, the direction and significance of main effects and interactions involving memory type, stimulus dimension, and search conditions were unaffected by inclusion of the target-side agreement variable; additionally, target-side agreement did not have significant interactions with any other variable included in the model.

We also conducted an analysis in which we added, in addition to the target-side agreement variable, a lure-side agreement variable. This lure-side agreement variable coded for whether the accessory item appeared on the same side as that indicated by the prompt or not (effect-coded: 1 = accessory feature on the same side as that indicated by the prompt; -1 = accessory feature on the opposite side as that indicated by the prompt). This analysis was only conducted for Inconsistent trials because it would not be applicable to the others: For Consistent trials, target-side agreement and lure-side agreement were necessarily the same, and there was no lure on Only trials. The model indicated no significant main effect of lure-side agreement, nor did this variable interact with any other predictors, including target-side agreement.

Because neither target-side agreement nor lure-side agreement variables interacted with any variables of interest, and inclusion of these variables did not affect our results, we report results from a subsequent model that omits these variables; this provides consistency with the models in prior experiments. Analyses reported are therefore identical to those conducted in Experiment 3.

Results

Training and Testing Phases

Participants successfully learned the association between each scene and its stimulus. In the Training Phase, mean (\pm SD) precision on the color/shapewheel (in degrees from the correct color/shape) generally improved over cycles: $8.74^\circ \pm 13.57^\circ$ for colors and $9.36^\circ \pm 15.79^\circ$ for shapes in cycle one; $7.05^\circ \pm 9.75^\circ$ for colors and $7.13^\circ \pm 11.71^\circ$ for shapes in cycle two; $6.57^\circ \pm 8.88^\circ$ for colors and $6.88^\circ \pm 12.42^\circ$ for shapes in cycle three.

The mean (\pm SD) number of cycles per scene needed to pass the Testing Phase was 1.44 ± 0.78 (for color-associated scenes, 1.48 ± 0.80 ; for shape-associated scenes, 1.41 ± 0.76), and 68.40% of scene-stimulus pairs only needed to be tested once (for color-associated scenes, 66.39%; for shape-associated scenes, 70.41%). Mean (\pm SD) precision on the color/shapewheel was $3.98^\circ \pm 2.63^\circ$ from the correct stimulus on the final test of each scene (i.e., when reported correctly from LTM): $4.12^\circ \pm 2.66^\circ$ for colors and $3.84^\circ \pm 2.60^\circ$ for shapes.

Search Phase

Search accuracy When searching for a stimulus feature in LTM, participants successfully clicked the prompted feature on the search display in $83.23\% \pm 11.21\%$ of trials, and when searching for a feature in WM, the rate was $95.72\% \pm 5.33\%$ of trials. When searching for a color, participants successfully clicked the prompted feature on the search display in $90.94\% \pm 7.50\%$ of trials, and when searching for a shape, the rate was

$88.00\% \pm 8.27\%$ of trials. No response was registered on the search display within the time allotted on $3.85\% \pm 3.31\%$ of LTM trials; on WM trials this rate was $1.02\% \pm 1.15\%$. On trials in which the prompted stimulus was a color, this rate was $1.99\% \pm 1.88\%$; and on trials with prompted shapes, this rate was $2.88\% \pm 2.43\%$.

Unprompted (color or shape) report Participants' unprompted reports after the search display were considered accurate ($\leq 20^\circ$ from the correct color or shape) on $83.08\% \pm 11.32\%$ of trials when reporting from LTM. On trials in which participants were reporting from WM, they were correct $82.94\% \pm 9.36\%$ of the time. When reporting a shape, $82.25\% \pm 10.47\%$ of trials were correct. When reporting a color, this rate was $83.77\% \pm 9.80\%$. When reporting the stimulus from LTM, $0.19\% \pm 0.68\%$ of trials did not receive a response in the time allotted; this never occurred when reporting from WM. When reporting a color, this rate was $0.21\% \pm 0.56\%$; and when reporting a shape, it was $0.12\% \pm 0.41\%$.

Search response times Response times and accuracy in the Search Phase are summarized in Table 1. As before, we only analyzed trials in which participants both (1) clicked on the correct stimulus feature in the search display and (2) reported the unprompted stimulus feature accurately. The mixed model analyzing log RT — as a function of memory type (WM vs. LTM), search condition (Consistent, Inconsistent, Only), and stimulus dimension (Shape, Color) — revealed, as in all previous Experiments, a main effect of memory type on log RTs such that participants were significantly slower to respond on the search display for LTM trials than for WM trials ($\beta = 0.073$, $SE = 0.005$, $p < 0.001$). See Fig. 9a for a plot of all fixed-effect estimates, and Fig. 9b for a plot of raw RT by memory type. In addition, a main effect of stimulus dimension was found, as in Experiment 3, indicating that participants were faster to respond on trials in which the prompted stimulus was a color rather than a shape ($\beta = 0.089$, $SE = 0.005$, $p < 0.001$).

As before, RTs were significantly faster on Consistent trials relative to Only trials ($\beta = -0.016$, $SE = 0.006$, $p = 0.004$; model-estimated RT benefit on Consistent vs. Only trials = 18.26 ms). There was also a significant effect of Inconsistent (vs. Only) search display ($\beta = 0.013$, $SE = 0.006$, $p = 0.026$; model-estimated RT cost = 14.69 ms), indicating that participants were slower to respond on Inconsistent versus Only trials.

As before, when the prompted stimulus feature was in the shape dimension (rather than a color), there was a larger effect of a trial being Consistent (relative to Only; $\beta = 0.014$, $SE = 0.006$, $p = 0.015$); and a larger effect of a trial being Inconsistent (relative to Only; $\beta = -0.029$, $SE = 0.006$, $p < 0.001$). That is, as in Experiment 3, when color was the accessory dimension (i.e., the trial had a prompted shape and an

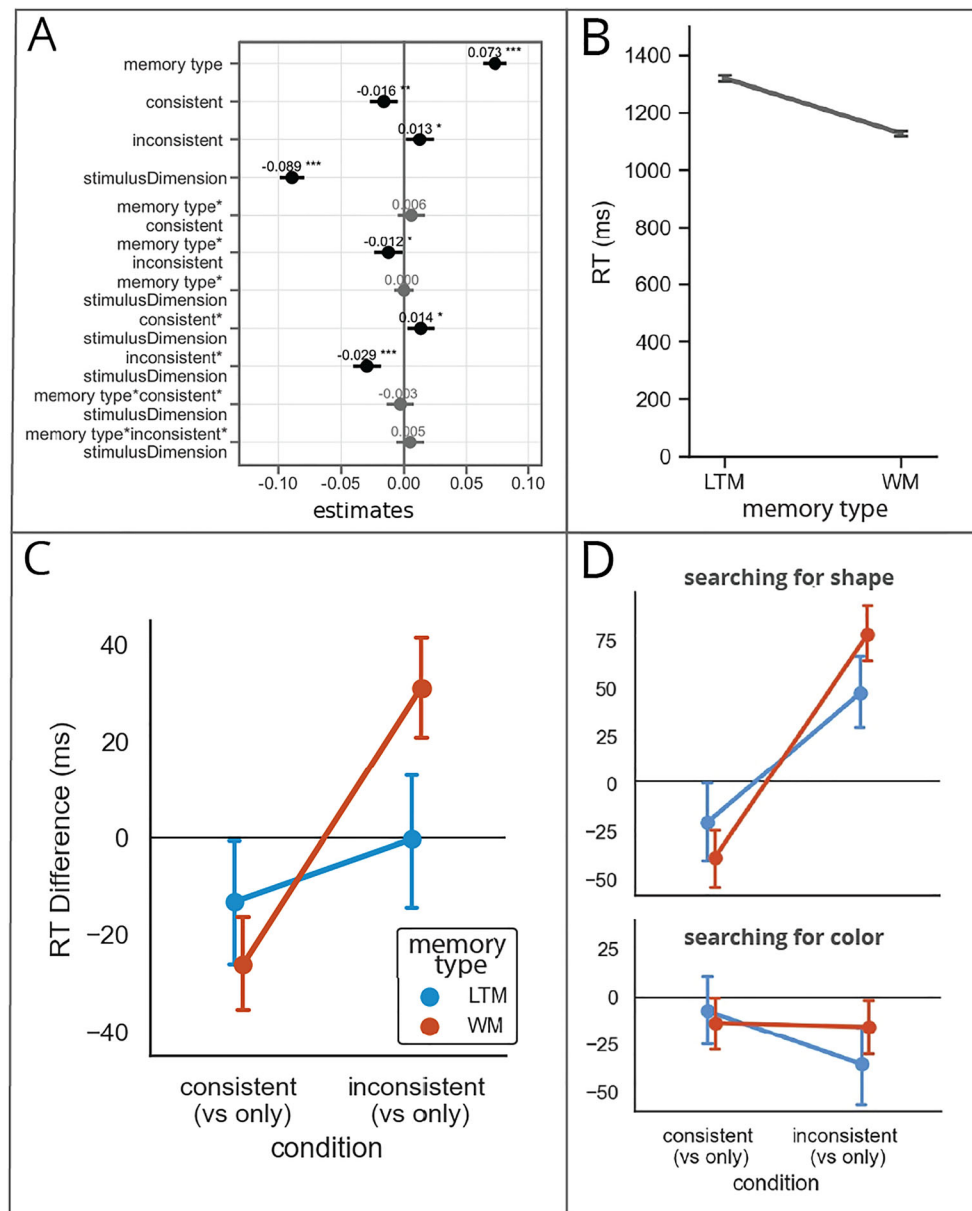


Fig. 9 Results from Experiment 4, in which participants were cued with either two long-term memory (LTM) representations or two working memory (WM) representations. **a** Coefficient estimates of the model for trial-wise log response time (RT). Error ribbons represent the 95% confidence interval; *** $p < .001$, ** $p < .01$, * $p < .05$. **b** LTM trials (vs. WM trials) were associated with a slower response on the search display. Error bars represent SEM of the within-participant (LTM-WM) difference. **c** Mean RT difference by memory type for Consistent (vs. Only) and Inconsistent (vs. Only) trials. Search RTs were facilitated on Consistent (vs. Only) search displays, and this facilitation did not interact with memory type. RTs were delayed on Inconsistent trials (vs. Only), but less so

when cued with two LTM cues than two WM cues. Error bars represent SEM of the within-participant difference between Consistent and Only and Inconsistent and Only trials. **d** Mean RT difference by dimension of the prompted stimulus for Consistent (vs. Only) and Inconsistent (vs. Only) trials. The model output – represented in (A) – confirms the interaction visualized here between stimulus dimension and search condition: The effects of a trial being Consistent (vs. Only) or Inconsistent (vs. Only) were larger for trials in which participants searched for a shape (and thus the unprompted stimulus was a color). Error bars represent SEM of the within-participant difference between Consistent and Only and Inconsistent and Only trials

unprompted color, rather than vice versa), the effects of having a Consistent or Inconsistent search display were greater.

Two differences emerged – relative to Experiment 3 – from the log RT model. Firstly, memory type no longer interacted with stimulus dimension, meaning that the RT benefit conferred by searching for a WM (vs. LTM) stimulus feature

was not different when that feature was a color vs. a shape ($\beta = < 0.001$, $SE = 0.004$, $p = 0.991$).

Critically, and unlike in Experiment 3, memory type and search condition interacted when the search condition was Inconsistent (relative to Only; $\beta = 0.012$, $SE = 0.006$, $p = 0.031$) such that the slowing of RT on Inconsistent trials was

greater on WM (vs. LTM) trials. There was no interaction between prompt and Consistent trials (relative to Only; $\beta = 0.006$, $SE = 0.006$, $p = 0.289$). To illustrate these effects, the RT data are plotted in milliseconds by prompt and condition in Fig. 9c, and by stimulus dimension and condition in Fig. 9d.

Comparison of search performance across Experiments 3 and 4

A follow-up linear mixed model on the log RT data was performed to compare Experiments 3 and 4. This comparison probes whether cooperation and/or competition during memory-guided search differ when the prompted and unprompted features are from different (Experiment 3) or the same (Experiment 4) memory source. If LTM and WM representations on a given trial are functionally the same, then no interaction should emerge that involves search condition and experiment version.

This trial-wise model was specified in the same way as the previous RT model, with the addition of an interacting fixed factor indicating experiment version (effect-coded: Experiment 3 = 1; Experiment 4 = -1). The consequent full experiment comparison model was: $\text{lmer}(\log(\text{RT}) \sim \text{memoryType} * \text{searchCondition} * \text{stimulusDimension} * \text{experiment} + (1 + \text{memoryType} * \text{stimulusDimension} | \text{Participant}))$. This random effects structure was specified on the same basis as before (see *Experiment 1/Methods/Data/Analysis*; by trimming the maximal structure variables by least explained variance using PCA until convergence).

Effects that were significant in each experiment individually were also significant in this combined model. We therefore observed the expected effects of memory type (LTM vs. WM), Consistent (vs. Only) search display, Inconsistent (vs. Only) search display, and stimulus dimension, as well as the finding that both search display condition effects (Consistent vs. Only and Inconsistent vs. Only) were larger when the unprompted feature was in the color dimension (see Fig. 10a).

A main effect of experiment version revealed that participants were faster to respond to the search display in Experiment 4 than in Experiment 3 ($\beta = 0.027$, $SE = 0.007$, $p < 0.001$). There was also a significant interaction between memory type and experiment such that the difference between WM trials and LTM trials was larger in Experiment 4 versus Experiment 3 ($\beta = -0.012$, $SE = 0.003$, $p < 0.001$). As anticipated, a significant three-way interaction between memory type, Inconsistent (vs. Only) search display, and experiment version ($\beta = 0.009$, $SE = 0.004$, $p = 0.038$) also emerged (Fig. 10b, right), reflecting the finding reported in the previous section (*Experiment 4/Results/Search Phase/Search response times*): Two LTM representations competed less than two WM representations in Experiment 4, while no asymmetry between LTM-prompted and WM-prompted trials was present in Experiment 3.

Lastly, this model revealed an interaction between experiment version and Consistent (vs. Only) search display such that the RT benefit on Consistent trials (relative to Only trials) was larger in Experiment 3 than in Experiment 4 ($\beta = -0.010$, $SE = 0.004$, $p = 0.017$). That is, when the prompted and unprompted memory representations were from different memory types (one LTM and one WM), participants showed more of a RT benefit on Consistent (vs. Only) trials, relative to when the two memories were of the same type (i.e., both WM or both LTM). This effect is visualized on the left side of Fig. 10b. The interactions observed between search performance and experiment are incompatible with the notion that all prompted memories are in the focus of attention, and all unprompted memories in the direct access region of WM, regardless of their source (WM or LTM). We discuss the implications of these results in the *General discussion*.

Discussion

In Experiment 3, we found no evidence that LTM cooperates or competes with WM more than the other way around. One potential explanation of these results is that LTM, when activated to guide attention, is placed into the same store as where WM information is placed to guide attention. That is, it is possible that no difference was observed as a function of memory type in Experiment 3 because the prompted memory may have been in the focus of attention regardless of its source (WM or LTM), and the unprompted memory may have been in the direct access region of WM, regardless of its source (Oberauer, 2002, 2009).

If it were indeed the case that cued LTMs become stored in the same way as cued WMs, then the same pattern of results (as Experiment 3) should hold if both memories on a given trial were accessed from WM, or both memories accessed from LTM. That is, regardless of the memory source, the prompted memory should be in the focus of attention and the unprompted memory in the direct access region of WM. To test this idea, we designed Experiment 4, which cued participants with either two LTM representations or two WM representations rather than one of each. Experiment 4 indicated that, similar to how LTM and WM representations cooperated and competed with one another to guide attention in Experiment 3, two representations of the same memory type also cooperated and competed. However, two key differences emerged.

The first key difference is that when two LTM representations could guide attention in the search task, less competition emerged than when attention could be guided by two WM representations. This led to a memory type (WM vs. LTM) by Inconsistent (vs. Only) by experiment (3 vs. 4) interaction, because there was no detectable difference in competition between WM-prompted (LTM accessory) and LTM-prompted (WM accessory) trials in Experiment 3.

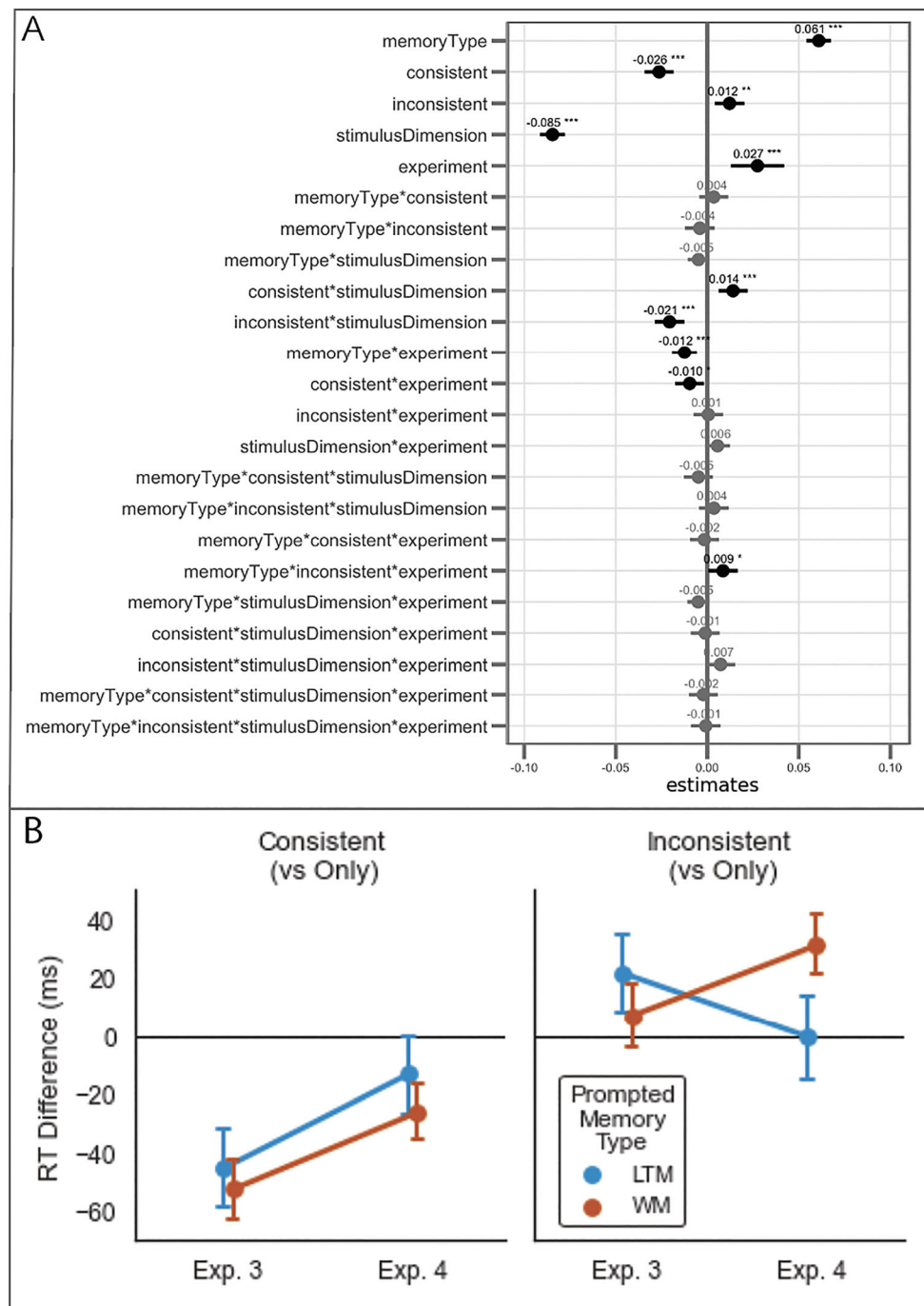


Fig. 10 Comparison of results from Experiments 3 and 4. In Experiment 3 the unprompted (accessory) feature was of the opposite memory type (i.e., if searching for a feature from long-term memory (LTM), the accessory feature was from working memory (WM)); in Experiment 4 the unprompted feature was of the same memory type (i.e., if searching for a feature from LTM, the accessory feature was also from LTM). **a** Coefficient estimates of the model for trial-wise log response time (RT) in the experiment comparison linear mixed model. Error ribbons represent the 95% confidence interval; *** $p < .001$, ** $p < .01$, * $p < .05$. **b** Mean RT difference by prompted memory type for Consistent (vs. Only)

and Inconsistent (vs. Only) trials across Experiments 3 and 4. The RT benefit on Consistent (vs. Only) trials was larger in Experiment 3 than Experiment 4 (Consistent (vs. Only) by experiment interaction), and the RT cost on Inconsistent (vs. Only) trials interacted with prompted memory type in Experiment 4 but not Experiment 3 (memory type by Inconsistent (vs. Only) by experiment interaction). Error bars represent SEM of the within-participant difference between Consistent and Only and Inconsistent and Only trials for each stimulus dimension and experiment

The second key difference is that two memories of the same type (i.e., either both LTM or both WM) cooperated less than two memories of different types (i.e., one from LTM and the other from WM). That is, cooperation was augmented when the memories came from different sources.

These findings are incompatible with the idea that LTM and WM, when activated to guide attention, are represented in a shared store, or otherwise identical format. We discuss the implications of our results in the *General discussion*, below

General discussion

Summary

Behavior is guided by memories at multiple timescales (Nobre & Stokes, 2019). The present study tested how working memory (WM) and long-term memory (LTM) guide visual attention when active in the same task. We found that WM and LTM representations readily cooperate during attentional guidance, and that they are also capable of competing under certain stimulus conditions. In particular, competition only occurred when individuals searched for a shape while attempting to avoid distraction from an irrelevant color. Across three studies, we found no evidence that WM cooperated or competed with LTM more than the other way around. A modified version of the task, however, in which two memories from the same source were cued (both WM or both LTM) revealed key differences between WM- and LTM-guided search: Two WM representations competed more than two LTM representations, and memories from different sources (WM and LTM) cooperated more than memories from the same source. Taken together, these results suggest that WM and LTM interact to guide visual search in the same task; that they compete and cooperate with each other in similar ways; but they are not represented in identical states during attentional guidance.

Proposed mechanisms

Cooperation between WM and LTM was observed in all experiments, but competition arose only when participants searched for a shape target and held an accessory color memory in mind. Why did we observe cooperation but not competition in Experiments 1 and 2, for which both memories were colors? The pattern of results we observed may be explained by an object-wise winner-takes-all mechanism that is active during memory-guided attention. We propose the following: First, visual features that match any activated memory representation (whether in WM or LTM, whether prioritized or not) contribute to evidence accumulation. Second, features that match prioritized (vs. accessory) memory representations lead to faster or stronger evidence accumulation (as reported by

Zhang et al., 2018). Third, some features (e.g., color, relative to shape) lead to faster or stronger evidence accumulation – perhaps because colors are typically more salient than shapes (Theeuwes, 1991). For example, color can elicit stronger attentional guidance effects than directional cues (Fan et al., 2021, Experiment 2) and irrelevant colors are more distracting than irrelevant shapes (Nickel et al., 2020; Theeuwes, 1991). Finally, evidence is summed over an object, and the object with the most accumulated evidence guides attention.

This model predicts that competition is stifled when the prioritized and accessory representations are both colors (as in Experiments 1 and 2), or both in the same stimulus dimension more generally. This is because, within the color dimension, features that match the prioritized (vs. accessory) memory representation led to faster evidence accumulation; without a second stimulus dimension, the object with the prioritized feature always “wins” the race to the decision threshold. This does not preclude *cooperation* between the prioritized and accessory memories: Because evidence is summed at the object level, features that match an accessory item can quicken evidence accumulation for that object.

When memories are in different stimulus dimensions, however, imbalance in the relative saliency of the visual features can lead to competition. Evidence from particularly salient features (e.g., color; Folk, 2015; Nickel et al., 2020) can rapidly accumulate. When the salient feature is a distractor, this evidence can accumulate faster than evidence for the less-salient but prioritized feature (e.g., a shape). This would produce more competition (relative to a non-salient distractor) when the distracting feature is in a different object than the prioritized memory, and more cooperation when it is in the same object. This model can therefore account for our observed effect of more competition, and more cooperation, when color is an accessory memory (during search for a shape) compared to the other way around.

The above model also needs to account for our finding that memories from different sources (WM and LTM) cooperate more than memories from the same source (two WMs or two LTMs). This may occur because of differential shielding between memories to prevent interference (similar to how WM can be shielded from perceptual distractors; see Hakim et al., 2020; Lorenc et al., 2021). Interference between items maintained in WM increases proportionately with the neural overlap that those item representations activate (Cohen et al., 2014; see also Yang et al., 2018; Oberauer & Lin, 2017). When two memory representations derive from the same (vs. different) source, they may engage more overlapping neural regions, and thus require more aggressive shielding to minimize the proportional increase in potential interference. Memories that are more efficiently shielded from one another may have less opportunity to cooperate to guide attention – thus accounting for why memories of the same source cooperate less than memories from different sources.

This notion of active-memory shielding may also help explain why we observed more competition between two WMs relative to two LTMs (Experiment 4). One possibility is that active LTMs may be more effectively shielded from one another than active WM representations. This would be consonant with the importance of pattern separation mechanisms for LTM – by which similar items come to be represented distinctly to reduce interference (Favila et al., 2016; Schurgin, 2018; Yassa & Stark, 2011; Zotow et al., 2020) – as well as research on efficient behavioral and neural suppression of potentially distracting LTMs (Anderson et al., 2004; Anderson & Green, 2001). Stronger shielding between activated LTM representations should lead to less competition between LTMs than between WMs, as evidenced in Experiment 4, and less cooperation between LTMs than between WMs (particularly when color is the accessory item, which, as noted above, is when cooperation and competition are most pronounced). While the latter effect – less cooperation for two LTMs versus two WMs – is not statistically significant in our data, the numerical trend is in that direction. Further work testing our proposed model will be important for determining if cooperation and competition within WM is reliably stronger than these dynamics within LTM.

There is, nevertheless, an alternative model that could explain cooperation between memories, in the form of performance enhancements on Consistent (vs. Only) trials. It is possible that, in Experiments 1 and 2, attention was only guided by the prompted color; however, once attention arrived at the appropriate object (half-circle), evidence accumulation was further enhanced (or a decision threshold dropped) once the accessory color was recognized to be in that same object (i.e., adjacent half-circle). Under this scenario, the accessory color does not guide attention but it does contribute to the visual search response either by enhancing further evidence accumulation once attention is already at the target object or by reducing the threshold for a response. With our procedures, we cannot tell the difference between these three possibilities in Experiments 1 and 2 (i.e., attentional guidance by the accessory color; enhanced evidence accumulation by the accessory color once attention is already guided to the target object; or a reduced decision threshold once the accessory color is detected to be adjacent to the target). Note, however, that whatever mechanism is at play does not affect the conclusion that working memory and long-term memory can cooperate to improve attentional behavior and visual search; it is merely the stage at which they cooperate that can be debated.

Importantly, we think that considering Experiments 1 and 2 together with Experiments 3 and 4 is more concordant with the hypothesis that multiple memories guide attention (as opposed to influencing evidence accumulation or decision thresholds once attention is already guided). In particular, it is not clear how competition between memories in guiding attention can arise if the effects of an accessory memory only

occur after attention has already been guided to the target item. That is, distraction or slowing on Inconsistent (vs. Only) trials is difficult to explain in terms of an altered decision threshold: Once attention is guided to the target feature, Only and Inconsistent trials are the same in the sense of having a distractor feature unitized with the target feature (i.e., a distractor shape or distractor color unitized with the target color or shape). Any slowing on Inconsistent (vs. Only) trials should therefore be due to attentional capture by the accessory item elsewhere in the display, rather than a different decision or evidence accumulation process that occurs after attention has already been guided to the target shape or color.

We believe it is parsimonious to explain competition and cooperation effects with the same mechanism (guidance by both activated memories, as we describe in our *Proposed mechanisms* above) and therefore prefer this interpretation. Alternatively, one can propose different mechanisms for cooperation and competition: That competition is due to attentional guidance by multiple memories, but cooperation is due to changes in decision thresholds or evidence accumulation once attention is already guided. These competing mechanisms can be tested in future work.

Finally, other factors not mentioned above may affect the object-wise winner-takes-all mechanism we propose. For example, the learned value of an item (Anderson et al., 2011) or its meaning (Henderson & Hayes, 2018) may increase the rate of evidence accumulation. Future studies that test these factors can be useful in extending this proposal of memory-guided attention.

Competition between working memory (WM) and long-term memory (LTM) in guiding attention

A main goal of the current investigations was to test whether WM and LTM compete symmetrically during attentional guidance, or if WM competes with LTM more than the other way around. Across Experiments 1–3, we failed to observe asymmetry between WM and LTM when guiding attention: There was no evidence that WM cooperated or competed with LTM more than vice versa. This lack of asymmetry cannot be explained by WM and LTM being functionally identical because, across all experiments, participants were reliably slower to search for a feature from LTM than one from WM. Although WM and LTM were not functionally identical, one possibility is that they came to be stored in a similar way, and that LTM may simply be weaker or slower to be stored in that state or format.

We therefore considered the possibility that the lack of asymmetry might be due to LTM being “placed in” WM to guide attention. If this is the case, then the prioritized memory may be represented in the focus of attention (FoA), regardless of its source (WM or LTM), and the accessory memory may be represented in the region of direct access (RDA) of working

memory, regardless of its source (WM or LTM) (Oberauer, 2002, 2009). If that were true, two LTMs should yield the same behavioral effects in attentional guidance as two WMs or one WM and one LTM.

Our results, however, do not accord with this strong interpretation of the concentric activation model. First, we found an interaction between memory type (WM vs. LTM) and Inconsistent (relative to Only) trials in Experiment 4: Two WM representations competed more than two LTM representations. This pattern of results was significantly different from that in Experiment 3, in which RT slowing on Inconsistent (vs. Only) trials was not different between LTM-prompted trials (WM accessory) and WM-prompted trials (LTM accessory). This result is contrary to the notion that the prioritized memory is always in the FoA and the accessory item always in the RDA regardless of their source.

Furthermore, we found an interaction between Experiment (3 vs. 4) and Consistent (vs. Only) trials, such that memories from different sources (WM and LTM) cooperate more than memories from the same source (both WM or both LTM). This is, again, incompatible with the idea that memories are represented similarly regardless of source (LTM or WM) and distinguished only based on their level of priority. Instead, this suggests that there are integral differences between these active memories' representational states.

Nevertheless, there may be a complex interaction between memory source (LTM or WM) and priority state (FoA, RDA, or otherwise), which could provide a potential amendment to the concentric activation model (Oberauer, 2002, 2009) that incorporates memory source. For example, if two LTMs are cued, the accessory LTM may remain in activated LTM (rather than the RDA) while the prioritized LTM is in the FoA. For the remaining cases we investigated (two WMs, or one WM and one LTM), the prioritized memory may be in the FoA and the accessory item in the RDA. In other words, if two LTMs are activated, the second (accessory) LTM may be more weakly represented than an accessory WM or a single accessory LTM. Such a scenario may explain why two WM representations compete more than two LTM representations: Two WM items are represented more similarly and more strongly than two LTM items. Further studies exploring these possibilities may be useful in elucidating asymmetries in how active LTMs and WMs are stored similarly versus differently, and the implications of that for how they interact during memory-guided attention.

One particularly interesting observation in our studies is that accessory LTMs consistently affected WM-guided search, despite WM-guided search being significantly faster than LTM-guided search. That is, LTM influenced visual search during trials for which search was much faster, on average, than LTM could yield on its own. This is consistent with a two-stage model of long-term memory retrieval (Ciaramelli et al., 2009; Moscovitch, 2008) in which an initial,

rapid retrieval process can guide behavior automatically, while a second, slower stage is necessary for conscious and deliberate access of those memories.

Relation to prior work

An ongoing debate in the memory-guided attention literature is whether only a single item, as opposed to multiple items, can guide attention in a given task. Our results align with the multi-item template hypothesis (Beck et al., 2012; Beck & Hollingworth, 2017): Across four experiments, we found that an accessory memory reliably sped visual search when its feature was contained in the same object as that of the prioritized memory. We observed that accessory memory features can also, under certain stimulus conditions, reliably slow visual search when they are in a different object than the prioritized memory features. These effects occurred despite accessory items conferring no advantage overall in the search task (most of the time, these features were either not present or could only hurt performance). Furthermore, we extend prior work in this domain – which typically focuses on multiple representations from WM – by showing that representations sourced from LTM, and memories from different sources, can guide attention in the same task as well.

This literature on multi-item memory-guided attention has tended to focus on competitive (distracting) effects during memory-guided attention (e.g., Bahle et al., 2018; Chen & Du, 2017; Fan et al., 2019; Frătescu et al., 2019; Hollingworth & Beck, 2016; van Moorselaar et al., 2014), though some also look at cooperative effects (e.g., Bahle et al., 2020). While a small selection of studies has looked at cooperation and competition in the same paradigm, (e.g., Fan et al., 2021, Experiment 1; Soto et al., 2005), no studies to our knowledge have reported the reliable imbalance between cooperative and competitive interactions exhibited in Experiments 1 and 2: We report that there is, under certain conditions, a bias towards cooperation (or, potentially, lack of competition altogether). Because there are such limited data regarding cooperative and competitive interactions in the same paradigm, however, we cannot determine precisely under which conditions this imbalance is exhibited. Further research will be needed to determine how and when biases towards cooperation (rather than competition) arise in memory-guided attention. Such research can determine, for example, if this imbalance is a feature of WM-LTM interactions during attentional guidance, or if it extends to guidance from two WMs or two LTMs as well.

Our work also adds to the literature on memory-guided attention by examining *bidirectional* cooperative and competitive interactions between WM and LTM. Some studies have looked at how WM may compete with LTM-guided search but did not explore the reverse relationship (Günseli et al., 2016). Other studies have examined how selection history (a

type of LTM) and contents in WM can jointly bias search (Schwark et al., 2013), or how accessory LTM features can capture attention during WM tasks (Fan & Turk-Browne, 2016, Experiment 2), but these studies lacked baselines to determine if the effects were due to cooperation, competition, or both. They were therefore unable to determine whether there are asymmetries in how WM and LTM cooperate or compete with each other in their respective paradigms (nor whether imbalances exist between competition and cooperation). We addressed these limitations in our study, finding evidence for both cooperation and competition between WM and LTM during memory-guided search. Furthermore, by comparing WM-LTM interactions to WM-WM and LTM-LTM interactions, we were able to find some differences in cooperative and competitive dynamics as a function of memory source. Thus, our work adds to prior literature on memory-guided attention by exploring cooperative and competitive interactions within the same task, and systematically comparing these effects based on memory source.

Limitations and future directions

Prior work proposing a “flexible gate” that mediates information transfer from LTM to WM (Oberauer, 2002, 2009) has focused on when LTM may interfere (or not) with content in WM. In this research, there was no demand to use this information to guide attention: The goal was simply to report information in memory. Thus, it is possible that we did not observe the predicted asymmetry (such that WM is relatively shielded from interfering LTM) because this gate’s role is primarily for protecting WM from proactive interference, and functions differently or not at all in the case of memory-guided attention. Further research is needed to compare how this gate functions for memory reporting vs. memory-guided attention in the same task.

It may also be the case that asymmetry between LTM and WM in guiding attention – whether due to the flexible gate model described above or otherwise – could arise in the accuracy of visual search, rather than its speed. Accuracy was near ceiling for WM trials in the current studies, limiting our ability to probe such effects. This was by design, given *a priori* predictions for RTs, and training procedures that were meant to maximize accuracy. Future studies could bring performance off of ceiling, e.g., with limited presentation durations or more aggressive visual masks, to determine if asymmetries in cooperation and competition arise in the accuracy of memory-guided attention.

An additional limitation is that the lack of asymmetry in cooperation or competition between WM and LTM in Experiments 1–3 may be due to limitations of the employed methods, or otherwise too low power, to detect the most subtle differences between conditions. However, we think it is unlikely that our studies were poorly powered for several

reasons. First, we were able to detect many two-way and three-way interactions, and were able to identify asymmetry in competition between memory types in Experiment 4 (in which two WM representations competed more than two LTM representations). Thus, Experiment 3 – which is identical to Experiment 4 except for which memory types were paired together on a given trial – is unlikely to be too poorly powered to detect asymmetry in competition.

Second, we conducted an additional analysis in which we combined Experiment 1 and Experiment 2 in one mixed-effects model that was identical to the analyses in the individual experiments, but with the addition of an “Experiment” variable. This allowed us to determine if the lack of competition in Experiments 1 and 2 was due to low power; however, we still did not detect statistically significant competition effects (Inconsistent vs. Only trials) in this larger, more well-powered analysis.

Finally, we also re-ran all log RT models with a higher threshold for participant inclusion: 60% of trials “both-correct” rather than 50%. Although this reduces the number of trials in the analyses (because fewer participants are included), it should also reduce noise by excluding participants who had the fewest trials overall. The direction and significance of all effects reported vis-a-vis the log RT models remained the same following implementation of this higher participant-inclusion threshold, with the exception of one effect: The prompt (WM or LTM) by stimulus dimension (color or shape) interaction observed in Experiment 3 was no longer significant. (This effect is not central to the paper, as it does not involve the critical “search condition” variable.)

While we have made substantial efforts to ensure that the present experiments have sufficient power to detect even relatively small differences in RTs – and were able to detect many subtle RT differences – studies with larger sample sizes, more sensitive methods, and/or more trials per participant may be able to detect smaller effects in cooperation or competition between memory types, should those effects exist.

In our studies, LTM-prompted trials were consistently slower than WM-prompted trials. We incorporated memory source as a covariate in all analyses; thus, these RT differences were controlled for when examining other effects. Nevertheless, future studies may consider providing additional training for participants in the LTM associations, until their RTs for LTM-prompted search are similar to those for WM-prompted search. This would allow exploration of how LTM and WM compete or cooperate when they are both brought to mind as quickly and easily. (Doing so, however, may further imbalance other measures of performance between LTM and WM trials – e.g., more LTM training may further increase the accuracy benefit for reported LTM vs. WM in the post-search report of the unprompted feature).

In the same vein, there are also differences in accuracy and RT by prompted stimulus dimension (color or shape). We

incorporated stimulus dimension as a covariate in all analyses; thus, these RT differences between shape- and color-prompted trials were controlled for when examining other effects. These RT differences likely occurred because the novel shapes in the Validated Circular Shape (VCS) space (Li et al., 2020) were relatively more difficult to remember and detect than familiar colors. Nevertheless, it was important for us to use the VCS space to keep both dimensions (color and shape) continuous, relatively perceptually uniform, and roughly comparable in terms of task demands on the post-search unprompted feature report. Nevertheless, future studies may consider employing familiar shape stimuli (like squares, circles, triangles) to permit investigation of whether asymmetry in competition between color and shape are eliminated in these circumstances (note, however, that Nickel et al., 2020 observed greater attentional capture by colors vs. shapes even though they used familiar shapes).

Although we find evidence that WM and LTM both guide attention in the same task, an open question is whether they do so simultaneously or in alternation, albeit on a rapid timescale. For example, active memories may be represented in oscillatory dynamics of nested theta-gamma subcycles (Lisman & Idiart, 1995; Sauseng et al., 2009; Wolinski et al., 2018), such that different memories are represented at different timepoints of a given cycle. Indeed, recent evidence suggests that two concurrently held items in WM are prioritized in alternation, not simultaneously, via an ongoing theta oscillation (Pomper & Ansorge, 2021). Our study was not designed to tease apart truly simultaneous versus rapid sequential representation of memories on a given trial. Probing the temporal dynamics of when WM and LTM are available to guide attention, and whether they are simultaneously or successively available, will require future research employing EEG, MEG, or behavioral studies carefully crafted to detect oscillations in behavior (e.g., Dehaene, 1993; Kerrén et al., 2022; Pomper & Ansorge, 2021; ter Wal et al., 2021; VanRullen, 2016).

Conclusion

At any given moment, multiple memories may inform our behavior, whether acquired recently (WM) or some time ago (LTM). We have shown that WM and LTM can work jointly to guide attention in the same task. WM and LTM reliably cooperate to guide attention and compete in more limited circumstances. We found no evidence that WM helps (or hinders) LTM-guided search more than the other way around. By directly comparing cooperation and competition between two WMs and two LTMs, we further showed that they are not functionally identical during memory-guided attention. Taken together, these results suggest that WM and LTM, despite functional differences, both cooperate and compete to guide attention.

Acknowledgements The authors would like to thank Serra Favila, Monica Thieu, and the Alyssano group for valuable insights on this project and advice about analyses.

Author' contribution Z.H.: conceptualization, methodology, software, validation, formal analysis, investigation, data curation, writing – original draft preparation, writing – review and editing, visualization, project administration.

M.A.: conceptualization, methodology, validation, resources, writing – original draft preparation, writing – review and editing, supervision, project administration, funding acquisition.

Funding This work was funded by an NSF CAREER Award (BCS-1844241) to M.A.

Data availability Data, materials, and code are available at <https://osf.io/y35ek>.

Declarations

Competing interest The authors have no relevant financial or non-financial interests to disclose.

Ethics approval Ethics approval was obtained from the Columbia University Institutional Review Board.

Informed consent All participants provided informed consent.

Consent publication All authors provided consent for publication.

References

- Aagten-Murphy, D., & Bays, P. M. (2019). Functions of Memory Across Saccadic Eye Movements. In T. Hodgson (Ed.), *Processes of Visuospatial Attention and Working Memory* (pp. 155–183). Springer International Publishing. https://doi.org/10.1007/7854_2018_66
- Aly, M., & Turk-Browne, N. B. (2017). How Hippocampal Memory Shapes, and Is Shaped by, Attention. In D. E. Hannula & M. C. Duff (Eds.), *The Hippocampus from Cells to Systems* (pp. 369–403). Springer International Publishing. https://doi.org/10.1007/978-3-319-50406-3_12
- Anderson, M. C., & Green, C. (2001). Suppressing unwanted memories by executive control. *Nature*, 410(6826), 366–369. <https://doi.org/10.1038/35066572>
- Anderson, B. A., & Halpern, M. (2017). On the value-dependence of value-driven attentional capture. *Attention, Perception, & Psychophysics*, 79(4), 1001–1011. <https://doi.org/10.3758/s13414-017-1289-6>
- Anderson, M. C., Ochsner, K. N., Kuhl, B., Cooper, J., Robertson, E., Gabrieli, S. W., Glover, G. H., & Gabrieli, J. D. E. (2004). Neural systems underlying the suppression of unwanted memories. *Science*, 303(5655), 232–235. <https://doi.org/10.1126/science.1089504>
- Anderson, B. A., Laurent, P. A., & Yantis, S. (2011). Value-driven attentional capture. *Proceedings of the National Academy of Sciences of the United States of America*, 108(25), 10367–10371. <https://doi.org/10.1073/pnas.1104047108>

- Atkinson, R. C., & Shiffrin, R. M. (1968). Human memory: A proposed system and its control processes. In *The psychology of learning and motivation* (pp. 89–195). Academic Press. [https://doi.org/10.1016/S0079-7421\(08\)60422-3](https://doi.org/10.1016/S0079-7421(08)60422-3)
- Awh, E., Belopolsky, A. V., & Theeuwes, J. (2012). Top-down versus bottom-up attentional control: A failed theoretical dichotomy. *Trends in Cognitive Sciences*, 16(8), 437–443. <https://doi.org/10.1016/j.tics.2012.06.010>
- Baddeley, A. (2003). Working memory: Looking back and looking forward. *Nature Reviews Neuroscience*, 4(10), 829–839. <https://doi.org/10.1038/nrn1201>
- Bahle, B., Beck, V. M., & Hollingworth, A. (2018). The architecture of interaction between visual working memory and visual attention. *Journal of Experimental Psychology: Human Perception and Performance*, 44(7), 992–1011. <https://doi.org/10.1037/xhp0000509>
- Bahle, B., Thayer, D. D., Mordkoff, J. T., & Hollingworth, A. (2020). The Architecture of Working Memory: Features from Multiple Remembered Objects Produce Parallel, Coactive Guidance of Attention in Visual Search. *Journal of Experimental Psychology: General*, 149(5), 967–983. <https://doi.org/10.1037/xge0000694>
- Baker, D. H., Vilidaite, G., Lygo, F. A., Smith, A. K., Flack, T. R., Gouws, A. D., & Andrews, T. J. (2021). Power Contours: Optimising Sample Size and Precision in Experimental Psychology and Human Neuroscience. *Psychological Methods*, 26(3), 295–314. <https://doi.org/10.1037/met0000337>
- Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, 67, 1–48. <https://doi.org/10.18637/jss.v067.i01>
- Beck, V. M., & Hollingworth, A. (2017). Competition in saccade target selection reveals attentional guidance by simultaneously active working memory representations. *Journal of Experimental Psychology: Human Perception and Performance*, 43(2), 225–230. <https://doi.org/10.1037/xhp0000306>
- Beck, V. M., Hollingworth, A., & Luck, S. J. (2012). Simultaneous Control of Attention by Multiple Working Memory Representations. *Psychological Science*, 23(8), 887–898. <https://doi.org/10.1177/0956797612439068>
- Chen, Y., & Du, F. (2017). Two visual working memory representations simultaneously control attention. *Scientific Reports*, 7, 6107. <https://doi.org/10.1038/s41598-017-05865-1>
- Chen, D., & Hutchinson, J. B. (2019). What Is Memory-Guided Attention? How Past Experiences Shape Selective Visuospatial Attention in the Present. *Current Topics in Behavioral Neurosciences*, 41, 185–212. https://doi.org/10.1007/7854_2018_76
- Ciaramelli, E., Lin, O., & Moscovitch, M. (2009). Episodic memory for spatial context biases spatial attention. *Experimental Brain Research*, 192(3), 511–520. <https://doi.org/10.1007/s00221-008-1548-9>
- Cohen, M. A., Konkle, T., Rhee, J. Y., Nakayama, K., & Alvarez, G. A. (2014). Processing multiple visual objects is limited by overlap in neural channels. *Proceedings of the National Academy of Sciences*, 111(24), 8955–8960. <https://doi.org/10.1073/pnas.1317860111>
- Cowan, N. (1993). Activation, attention, and short-term memory. *Memory & Cognition*, 21(2), 162–167. <https://doi.org/10.3758/BF03202728>
- Cowan, N. (1998). Memory activation and attentional focusing. In *Attention and Memory*. Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780195119107.003.0003>
- Cowan, N. (2008). What are the differences between long-term, short-term, and working memory? *Progress in Brain Research*, 169, 323–338. [https://doi.org/10.1016/S0079-6123\(07\)00020-9](https://doi.org/10.1016/S0079-6123(07)00020-9)
- Danek, R. H., & Mordkoff, J. T. (2011). Unequal motor durations under simple-, go/no-go, and choice-RT tasks: Extension of Miller and Low (2001). *Journal of Experimental Psychology: Human Perception and Performance*, 37(4), 1323–1329. <https://doi.org/10.1037/a0023092>
- Dehaene, S. (1993). Temporal oscillations in human perception. *Psychological Science*, 4(4), 264–270. <https://doi.org/10.1111/j.1467-9280.1993.tb00273.x>
- Desimone, R., & Duncan, J. (1995). Neural mechanisms of selective visual attention. *Annual Review of Neuroscience*, 18(1), 193–222. <https://doi.org/10.1146/annurev.ne.18.030195.001205>
- Downing, P. E. (2000). Interactions between visual working memory and selective attention. *Psychological Science*, 11(6), 467–473. <https://doi.org/10.1111/1467-9280.00290>
- Fan, J. E., & Turk-Browne, N. B. (2016). Incidental biasing of attention from visual long-term memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 42(6), 970–977. <https://doi.org/10.1037/xlm0000209>
- Fan, L., Sun, M., Xu, M., Li, Z., Diao, L., & Zhang, X. (2019). Multiple representations in visual working memory simultaneously guide attention: The type of memory-matching representation matters. *Acta Psychologica*, 192, 126–137. <https://doi.org/10.1016/j.actpsy.2018.11.005>
- Fan, L., Xu, M., ul Hasan, N., Sun, M., & Zhang, X. (2021). Attentional guidance from two representations of the same or different dimensions in visual working memory. *Current Psychology*, 40(1), 355–366. <https://doi.org/10.1007/s12144-018-9949-9>
- Favila, S. E., Chanales, A. J., & Kuhl, B. A. (2016). Experience-dependent hippocampal pattern differentiation prevents interference during subsequent learning. *Nature Communications*, 7(1), 1–10. <https://doi.org/10.1038/ncomms11066>
- Folk, C. L. (2015). The role of color in the voluntary and involuntary guidance of selective attention. In A. J. Elliot, A. Franklin, & M. D. Fairchild (Eds.), *Handbook of Color Psychology* (pp. 481–501). Cambridge University Press. <https://doi.org/10.1017/CBO9781107337930.024>
- Frätescu, M., Van Moorselaar, D., & Mathôt, S. (2019). Can you have multiple attentional templates? Large-scale replications of Van Moorselaar, Theeuwes, and Olivers (2014) and Hollingworth and Beck (2016). *Attention, Perception, & Psychophysics*, 81(8), 2700–2709. <https://doi.org/10.3758/s13414-019-01791-8>
- Fuster, J. (1997). Network Memory. *Trends in Neurosciences*, 20, 451–459. [https://doi.org/10.1016/S0166-2236\(97\)01128-4](https://doi.org/10.1016/S0166-2236(97)01128-4)
- Günseli, E., Olivers, C. N. L., & Meeter, M. (2016). Task-irrelevant memories rapidly gain attentional control with learning. *Journal of Experimental Psychology: Human Perception and Performance*, 42(3), 354–362. <https://doi.org/10.1037/xhp0000134>
- Gutwin, C., Cockburn, A., & Coveney, A. (2017). Peripheral Popout: The Influence of Visual Angle and Stimulus Intensity on Popout Effects. Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems, 208–219. <https://doi.org/10.1145/3025453.3025984>
- Hakim, N., Feldmann-Wüstefeld, T., Awh, E., & Vogel, E. K. (2020). Perturbing Neural Representations of Working Memory with Task-irrelevant Interruption. *Journal of Cognitive Neuroscience*, 32(3), 558–569. https://doi.org/10.1162/jocn_a_01481
- Henderson, J. M., & Hayes, T. R. (2018). Meaning guides attention in real-world scene images: Evidence from eye movements and meaning maps. *Journal of Vision*, 18(6), 10. <https://doi.org/10.1167/18.6.10>
- Hollingworth, A., & Beck, V. M. (2016). Memory-based attention capture when multiple items are maintained in visual working memory. *Journal of Experimental Psychology: Human Perception and Performance*, 42(7), 911–917. <https://doi.org/10.1037/xhp0000230>
- Hutchinson, J. B., & Turk-Browne, N. B. (2012). Memory-guided attention: Control from multiple memory systems. *Trends in Cognitive Sciences*, 16(12), 576–579. <https://doi.org/10.1016/j.tics.2012.10.003>
- Kerrén, C., van Bree, S., Griffiths, B. J., & Wimber, M. (2022). Phase separation of competing memories along the human hippocampal theta rhythm. *BioRxiv*. <https://doi.org/10.1101/2022.05.07.490872>

- Koch, C., & Ullman, S. (1985). Shifts in selective visual attention: Towards the underlying neural circuitry. *Human Neurobiology*, 4(4), 219–227.
- Konkle, T., Brady, T. F., Alvarez, G. A., & Oliva, A. (2010). Scene memory is more detailed than you think: The role of categories in visual long-term memory. *Psychological Science*, 21(11), 1551–1556. <https://doi.org/10.1177/0956797610385359>
- Li, A. Y., Liang, J. C., Lee, A. C. H., & Barense, M. D. (2020). The validated circular shape space: Quantifying the visual similarity of shape. *Journal of Experimental Psychology: General*, 149(5), 949–966. <https://doi.org/10.1037/xge0000693>
- Lisman, J. E., & Idiart, M. A. (1995). Storage of 7 +/- 2 short-term memories in oscillatory subcycles. *Science*, 267(5203), 1512–1515. <https://doi.org/10.1126/science.7878473>
- Lo, S., & Andrews, S. (2015). To transform or not to transform: Using generalized linear mixed models to analyse reaction time data. *Frontiers in Psychology*, 6, 1171. <https://doi.org/10.3389/fpsyg.2015.01171>
- Lorenc, E. S., Mallett, R., & Lewis-Peacock, J. A. (2021). Distraction in Visual Working Memory: Resistance is Not Futile. *Trends in Cognitive Sciences*, 25(3), 228–239. <https://doi.org/10.1016/j.tics.2020.12.004>
- Lüdtke, D. (2018). sjPlot: Data visualization for statistics in social science. *R Package Version*, 2(1).
- Miller, J. (1982). Divided attention: Evidence for coactivation with redundant signals. *Cognitive Psychology*, 14(2), 247–279. [https://doi.org/10.1016/0010-0285\(82\)90010-X](https://doi.org/10.1016/0010-0285(82)90010-X)
- Miller, J. O., & Low, K. (2001). Motor processes in simple, go/no-go, and choice reaction time tasks: A psychophysiological analysis. *Journal of Experimental Psychology: Human Perception and Performance*, 27(2), 266–289.
- Mizrak, E., & Oberauer, K. (2021). Working memory recruits long-term memory when it is beneficial: Evidence from the Hebb effect. *Journal of Experimental Psychology: General*. <https://doi.org/10.1037/xge0000934>
- Mordkoff, J., & Yantis, S. (1993). Dividing attention between color and shape: Evidence of coactivation. *Perception & Psychophysics*. <https://doi.org/10.3758/BF03206778>
- Moscovitch, M. (2008). The hippocampus as a “stupid”, domain-specific module: Implications for theories of recent and remote memory, and of imagination. *Canadian Journal of Experimental Psychology/Revue Canadienne de Psychologie Expérimentale*, 62(1), 62–79. <https://doi.org/10.1037/1196-1961.62.1.62>
- Nickel, A. E., Hopkins, L. S., Minor, G. N., & Hannula, D. E. (2020). Attention capture by episodic long-term memory. *Cognition*, 201, 104312. <https://doi.org/10.1016/j.cognition.2020.104312>
- Nobre, A. C., & Stokes, M. G. (2019). Premembering Experience: A Hierarchy of Time-Scales for Proactive Attention. *Neuron*, 104(1), 132–146. <https://doi.org/10.1016/j.neuron.2019.08.030>
- Oberauer, K. (2002). Access to information in working memory: Exploring the focus of attention. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 28(3), 411. <https://doi.org/10.1037/0278-7393.28.3.411>
- Oberauer, K. (2009). Chapter 2 design for a working memory. In *Psychology of learning and motivation* (Vol. 51, pp. 45–100). Academic Press. [https://doi.org/10.1016/S0079-7421\(09\)51002-X](https://doi.org/10.1016/S0079-7421(09)51002-X)
- Oberauer, K., & Lin, H.-Y. (2017). An interference model of visual working memory. *Psychological Review*, 124(1), 21–59. <https://doi.org/10.1037/rev0000044>
- Oberauer, K., Awh, E., & Sutterer, D. W. (2017). The role of long-term memory in a test of visual working memory: Proactive facilitation but no proactive interference. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 43(1), 1–22. <https://doi.org/10.1037/xlm0000302>
- Pomper, U., & Ansorge, U. (2021). Theta-Rhythmic Oscillation of Working Memory Performance. *Psychological Science*, 32(11), 1801–1810. <https://doi.org/10.1177/09567976211013045>
- R Core Team. (2022). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing.
- Ranganath, C., & Blumenfeld, R. S. (2005). Doubts about double dissociations between short- and long-term memory. *Trends in Cognitive Sciences*, 9(8), 374–380. <https://doi.org/10.1016/j.tics.2005.06.009>
- Sauseng, P., Klimesch, W., Heise, K. F., Gruber, W. R., Holz, E., Karim, A. A., Glennon, M., Gerloff, C., Birbaumer, N., & Hummel, F. C. (2009). Brain oscillatory substrates of visual short-term memory capacity. *Current Biology*, 19(21), 1846–1852. <https://doi.org/10.1016/j.cub.2009.08.062>
- Schurgin, M. W. (2018). Visual memory, the long and the short of it: A review of visual working memory and long-term memory. *Attention, Perception, & Psychophysics*, 80(5), 1035–1056. <https://doi.org/10.3758/s13414-018-1522-y>
- Schwark, J. D., Dolgov, I., Sandry, J., & Volkman, C. B. (2013). Simultaneous attentional guidance by working-memory and selection history reveals two distinct sources of attention. *Acta Psychologica*, 144(2), 269–278. <https://doi.org/10.1016/j.actpsy.2013.06.017>
- Shin, H., & Ma, W. J. (2016). Crowdsourced single-trial probes of visual working memory for irrelevant features. *Journal of Vision*, 16(5), 10. <https://doi.org/10.1167/16.5.10>
- Soto, D., Heinke, D., Humphreys, G. W., & Blanco, M. J. (2005). Early, involuntary top-down guidance of attention from working memory. *Journal of Experimental Psychology: Human Perception and Performance*, 31(2), 248–261. <https://doi.org/10.1037/0096-1523.31.2.248>
- Soto, D., Hodsoll, J., Rotshtein, P., & Humphreys, G. W. (2008). Automatic guidance of attention from working memory. *Trends in Cognitive Sciences*, 12(9), 342–348. <https://doi.org/10.1016/j.tics.2008.05.007>
- Stokes, M. G., Atherton, K., Patai, E. Z., & Nobre, A. C. (2012). Long-term memory prepares neural activity for perception. *Proceedings of the National Academy of Sciences*, 109(6), E360–E367. <https://doi.org/10.1073/pnas.1108555108>
- Summerfield, J. J., Lepsien, J., Gitelman, D. R., Mesulam, M. M., & Nobre, A. C. (2006). Orienting Attention Based on Long-Term Memory Experience. *Neuron*, 49(6), 905–916. <https://doi.org/10.1016/j.neuron.2006.01.021>
- ter Wal, M., Linde-Domingo, J., Lifanov, J., Roux, F., Kolibius, L. D., Gollwitzer, S., Lang, J., Hamer, H., Rollings, D., Sawlani, V., Chelvarajah, R., Staresina, B., Hanslmayr, S., & Wimber, M. (2021). Theta rhythmicity governs human behavior and hippocampal signals during memory-dependent tasks. *Nature Communications*, 12(1), 7048. <https://doi.org/10.1038/s41467-021-27323-3>
- Theeuwes, J. (1991). Cross-dimensional perceptual selectivity. *Perception & Psychophysics*, 50(2), 184–193. <https://doi.org/10.3758/BF03212219>
- Treisman, A. M., & Gelade, G. (1980). A feature-integration theory of attention. *Cognitive Psychology*, 12(1), 97–136. [https://doi.org/10.1016/0010-0285\(80\)90005-5](https://doi.org/10.1016/0010-0285(80)90005-5)
- van Ede, F., & Nobre, A. C. (2022). A neural decision signal during internal sampling from working memory. *BioRxiv*. <https://doi.org/10.1101/2022.03.31.486618>
- van Ede, F., & Nobre, A. C. (2023). Turning attention inside out: How working memory serves behavior. *Annual Review of Psychology, Forthcoming*. <https://doi.org/10.1146/annurev-psych-021422-041757>
- van Moorselaar, D., Theeuwes, J., & Olivers, C. N. L. (2014). In competition for the attentional template: Can multiple items within visual working memory guide attention? *Journal of Experimental*

- Psychology. *Human Perception and Performance*, 40(4), 1450–1464. <https://doi.org/10.1037/a0036229>
- VanRullen, R. (2016). Perceptual Cycles. *Trends in Cognitive Sciences*, 20(10), 723–735. <https://doi.org/10.1016/j.tics.2016.07.006>
- Verschooren, S., Kessler, Y., & Egner, T. (2021). Evidence for a single mechanism gating perceptual and long-term memory information into working memory. *Cognition*, 212, 104668. <https://doi.org/10.1016/j.cognition.2021.104668>
- Vo, V. A., Sutterer, D. W., Foster, J. J., Sprague, T. C., Awh, E., & Serences, J. T. (2022). Shared Representational Formats for Information Maintained in Working Memory and Information Retrieved from Long-Term Memory. *Cerebral Cortex*, 32(5), 1077–1092. <https://doi.org/10.1093/cercor/bhab267>
- Wickham, H. (2018). *ggplot2: Elegant graphics for data analysis* (2nd ed.). Springer.
- Wolinski, N., Cooper, N. R., Sauseng, P., & Romei, V. (2018). The speed of parietal theta frequency drives visuospatial working memory capacity. *PLoS Biology*, 16(3), e2005348. <https://doi.org/10.1371/journal.pbio.2005348>
- Yang, L., Mo, L., Wang, X., & Yu, M. (2018). The effects of categorical similarity and feature-space proximity on visual working memory processing. *Visual Cognition*, 26(2), 100–114. <https://doi.org/10.1080/13506285.2017.1390515>
- Yassa, M. A., & Stark, C. E. L. (2011). Pattern separation in the hippocampus. *Trends in Neurosciences*, 34(10), 515–525. <https://doi.org/10.1016/j.tins.2011.06.006>
- Zhang, B., Liu, S., Doro, M., & Galfano, G. (2018). Attentional Guidance from Multiple Working Memory Representations: Evidence from Eye Movements. *Scientific Reports*, 8(1), 13876. <https://doi.org/10.1038/s41598-018-32144-4>
- Zhou, C., Lorient, M. M., & Mathôt, S. (2020). Concurrent guidance of attention by multiple working memory items: Behavioral and computational evidence. *Attention, Perception, & Psychophysics*, 82(6), 2950–2962. <https://doi.org/10.3758/s13414-020-02048-5>
- Zotow, E., Bisby, J. A., & Burgess, N. (2020). Behavioral evidence for pattern separation in human episodic memory. *Learning & Memory*, 27(8), 301–309. <https://doi.org/10.1101/lm.051821.120>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.