The effect of memory load on object reconstruction: Insights from an online mouse-tracking task

Aedan Y. Li[1] · James Y. Yuan[1] · Carson Pun[1] · Morgan D. Barense[1]

Abstract

Why can't we remember everything that we experience? Previous work in the domain of object memory has suggested that our ability to resolve interference between relevant and irrelevant object features may limit how much we can remember at any given moment. Here, we developed an online mouse-tracking task to study how memory load influences object reconstruction, testing participants synchronously over virtual conference calls. We first tested up to 18 participants concurrently, replicating memory findings from a condition where participants were tested individually. Next, we examined how memory load influenced mouse trajectories as participants reconstructed target objects. We found interference between the contents of working memory and what was perceived during object reconstruction, an effect that interacted with visual similarity and memory load. Furthermore, we found interference from previously studied but currently irrelevant objects, providing evidence of object-to-location binding errors. At the greatest memory load, participants were nearly three times more likely to move their mouse cursor over previously studied nontarget objects, an effect observed primarily during object reconstruction rather than in the period before the final response. As evidence of the dynamic interplay between working memory and perception, these results show that object reconstruction behavior may be altered by (i) interference between what is represented in mind and what is currently being viewed, and (ii) interference from previously studied but currently irrelevant information. Finally, we discuss how mouse tracking can provide a rich characterization of participant behavior at millisecond temporal resolution, enormously increasing power in cognitive psychology experiments.

Keywords Mouse tracking · Interference · Object-to-location binding · Online executable files · Virtual conference calls · Simultaneous reconstruction task

The capacity of visual working memory is surprisingly limited. Although we encounter thousands of objects during our daily lives, we can remember only a few pieces of information at any moment (Cowan, 2001, 2010). Some research has suggested that this capacity limit is closely related to our ability to resolve the interference from competing information (Bartsch & Oberauer, 2023; Endress & Szabó, 2017; Oberauer & Lin, 2017; Shipstead & Engle, 2013). For example, to identify a lemon, we must correctly associate its shape ("oval"), color ("yellow"), and location ("on the table") while distinguishing the lemon from objects with similar features, such as a lime or an apple. This interference from overlapping features can be minimized by forming distinct object representations through *feature binding* (Cowell et al., 2019; Hedayati et al., 2022; A. Y. Li et al., 2022; Oberauer, 2019; Schneegans & Bays, 2017; Swan & Wyble, 2014). When feature binding fails, *binding errors* can occur (Treisman, 1996), where between-item interference can result in the incorrect combination of features belonging to different objects. These errors manifest commonly in everyday life, such as misremembering the proper fruit for a dinner recipe but also in more severe conditions, such as Alzheimer's disease, which has been associated with deficits in visual binding (Parra et al., 2009).

✉ Aedan Y. Li
aedanyue.li@utoronto.ca

✉ James Y. Yuan
jy.yuan@mail.utoronto.ca

[1] Department of Psychology, University of Toronto, 100 St. George Street, Toronto, ON M5S 3G3, Canada

The study of feature binding in working memory has a rich history in the psychological sciences. Early influential models considered binding to operate over independent feature maps, with attention serving as the "glue" between features sharing common spatial locations (Treisman & Gelade, 1980). This seminal work led to debates about the format of the object representation, with experiments positing shape and color features can be bound directly to each other (Luck & Vogel, 1997; Zhang & Luck, 2008), bound by virtue of space or time (Schneegans et al., 2022; Schneegans & Bays, 2017), or bound hierarchically as both integrated wholes and independent features (Hedayati et al., 2022; A. Y. Li et al., 2022). Converging findings reveal that the format of the object representation can likely include multiple forms of binding, such that features may be initially bound to shared spatial maps but are then bound to form object representations that no longer depend on space (Shepherdson et al., 2022). This view has support from neuroimaging research whereby early posterior neocortex may represent features bound to location (Henderson et al., 2022; Hubel & Wiesel, 1962; M. Li et al., 2014; Schneegans & Bays, 2017; Sprague & Serences, 2013; Thyer et al., 2022), with fully specified objects bound to spatial context represented in anterior regions like the medial temporal lobes (Cooper & Ritchey, 2019; Cowell et al., 2019; Liang et al., 2020; Martin et al., 2018; Wu & Buckley, 2022; Yeung et al., 2013; Yeung et al., 2017; Yeung et al., 2019; Yonelinas et al., 2019). Critically, this body of work suggests a division between feature-to-location binding that may depend on posterior regions of the neocortex with object-to-location binding that may depend on anterior regions such as the medial temporal lobes.

To study feature binding, working memory researchers often quantify the errors that occur when binding fails. For example, one influential procedure quantifies feature-to-location binding errors using continuous reconstruction tasks (Ma et al., 2014). In a typical variant of this task, a number of colored squares are first shown during the study phase of an experiment. After a delay period, participants are then cued to reconstruct the color that was studied at an indicated location. Participants select the target color by moving their mouse cursor along a circular color wheel. Studies like these quantify binding errors as the proportion of participants' final selections that correspond with the uncued, nontarget colors, which have been incorrectly associated with the target location. When estimated using mixture models, these nontarget feature responses are known as "swap" errors (see Bays, 2016; Bays et al., 2009). This work typically finds that feature-to-location binding errors increase in tandem with memory load, suggesting that failures in feature integration are more likely to occur as participants attempt to hold increasingly more information in mind. However, while existing research in working memory has most often studied binding errors using simpler features like color or orientation when cued by location (i.e.,

feature-to-location binding errors), it remains an open question how memory load may influence binding errors for more complex objects during working memory (i.e., object-to-location binding errors).

One previous method to study complex feature binding errors has been through eye tracking (Barense et al., 2012; Erez et al., 2013; Ryan et al., 2007; Yeung et al., 2013; Yeung et al., 2017; Yeung et al., 2019), which provides rich information about the internal representation at millisecond temporal resolution (for reviews, see Hannula et al., 2010; Kragel & Voss, 2022; Ryan et al., 2020; Voss et al., 2017; Wynn et al., 2019). For example, object-to-location binding errors can be measured as the proportion of eye fixations made towards the relevant target (i.e., previously studied objects) compared with irrelevant nontarget objects (i.e., similar lures) over the entire gaze trajectory. This body of work has found that patients with medial temporal lobe damage (Barense et al., 2012; Erez et al., 2013; Ryan et al., 2000) and older adults at risk for Alzheimer's disease (Yeung et al., 2013; Yeung et al., 2017; Yeung et al., 2019) have aberrant viewing behavior compared with healthy older adults. Populations with medial temporal lobe damage often fail to direct their gaze towards important locations in scenes, evidence that these regions may be essential for successfully binding complex objects to locations within spatial environments (Ryan et al., 2000; Yeung et al., 2019). Analyzing the continuous trajectory of eye movements can therefore be used to quantify object-to-location binding errors, operationalized as the fixations made towards the location of irrelevant objects in a spatial environment.

The majority of existing working memory studies quantify binding errors using only the final participant response during the test phase of an experiment (Fig. 1a) without considering the entire trajectory of navigating to those responses (Fig. 1b; but see Hao et al., 2021; Park & Zhang, 2022, for recent examples of mouse tracking in a working memory experiment). The standard approach may exclude important information, akin to analyzing only the final gaze position during eye tracking without considering the eye movements that form the path leading up to the final gaze position (see Barense et al., 2012; Erez et al., 2013; Golomb et al., 2008; Golomb & Kanwisher, 2012; Golomb et al., 2014; Hannula et al., 2010; Liu et al., 2017; Wynn et al., 2020; Yeung et al., 2017; Yeung et al., 2019, for examples of eye-tracking experiments that consider the scan path across the entire trial). Furthermore, existing working memory research has primarily focused on simpler feature-to-location binding errors, limiting our understanding of how binding operates over more complex object features in working memory. Inspired by eye-tracking methodologies, we quantified object-to-location binding errors in the present study not only from the final response at test but also as participants continuously reconstructed shape–color objects
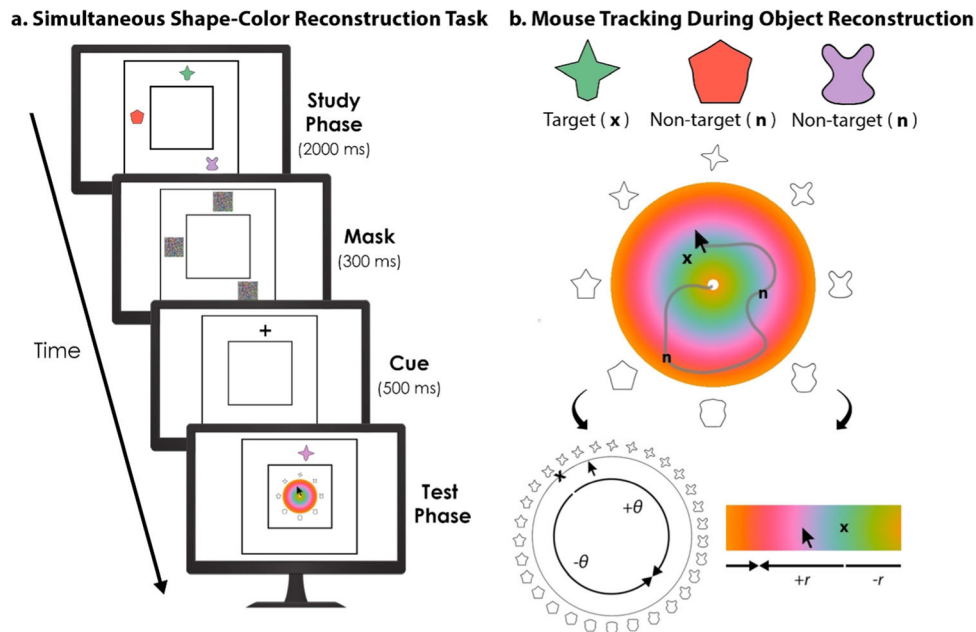
Fig. 1 **a** Simultaneous shape–color reconstruction task. To ensure that the task could be displayed on most participant computers, the task appeared in a box spanning 1,080 × 1,080 pixels. Participants were asked to remember shape–color objects during the study phase for 2,000 ms (set size was manipulated according to Fig. 3b). After an ISI of 300 ms, a mask appeared at the positions of the studied objects (300 ms). There was then a retention interval of 1,000 ms between the masks and the cue where no items were shown. The location of one study object was then cued by a fixation cross (500 ms), and the test phase display appeared with the mouse cursor positioned at the centre of the screen. Participants reconstructed the cued, target shape and color along a 2D response space (untimed; shown enlarged in panel **b**; for more details, see A. Y. Li et al., 2022). While the participant was reconstructing the target object, the program displayed their reconstruction corresponding to their mouse cursor position on shape and color space at the cued location (see video example: https://osf.io/ycq5s). Thus, participants were required to match the target object from memory with what is perceived on the display using the mouse cursor. Critically, along the circle's circumference, we mapped the Validated Circular Shape Space (A. Y. Li et al., 2020), and along the radius we mapped a circular color wheel sampled from CIELAB color space. **b** We recorded the participant mouse movement (gray line) continuously during the entire test phase as participants reconstructed objects. In this way, we could quantify the mouse movements made towards target and nontarget object features dynamically during object reconstruction. On the above, θ refers to angular distance from the target shape, whereas r refers to radial distance from the target color. (Color figure online)

using mouse tracking. This approach enabled us to convert a dependent variable with a single observation each trial (i.e., the final response on a cognitive task) to a dependent variable with hundreds of observations each trial (i.e., time-series data at millisecond temporal resolution), providing a rich characterization of participant behavior leading up to the final response at test.

To test object-to-location binding errors in working memory, we adapted a simultaneous reconstruction task to include online mouse tracking (Fig. 1b; A.Y. Li et al., 2022). Drawing from previous eye-tracking literature, we define object-to-location binding errors as the between-item interference stemming from irrelevant lure objects from study. We recorded the mouse path trajectories every trial as participants reconstructed objects (analogous to recording gaze trajectories from an eye-tracking experiment). Moreover, we developed our task as a downloadable executable file so that participants could access the task online and run it on their own machines, and we tested the data reliability and testing efficiency of this experimental method when paired with virtual conferencing

rooms. Thus, in addition to studying the effect of memory load on object-to-location binding errors for more complex shape–color objects than commonly tested in the literature (see Ma et al., 2014), this work provides a novel online approach to quantify mouse-path trajectories while participants reconstruct objects from memory.

Method

We first examined the *data reliability* of online executables when paired with virtual conference rooms (Fig. 2), attempting to replicate previous memory load findings from the literature (A. Y. Li et al., 2022; Ma et al., 2014; Sone et al., 2021). In the individual testing condition, participants downloaded an executable program on their own computers, and then completed the memory task (Fig. 3a) one-on-one with the experimenter in virtual conference calls (i.e., over Zoom; Fig. 3b). Critically, we also examined the *testing efficiency* of our virtual conference room approach, testing
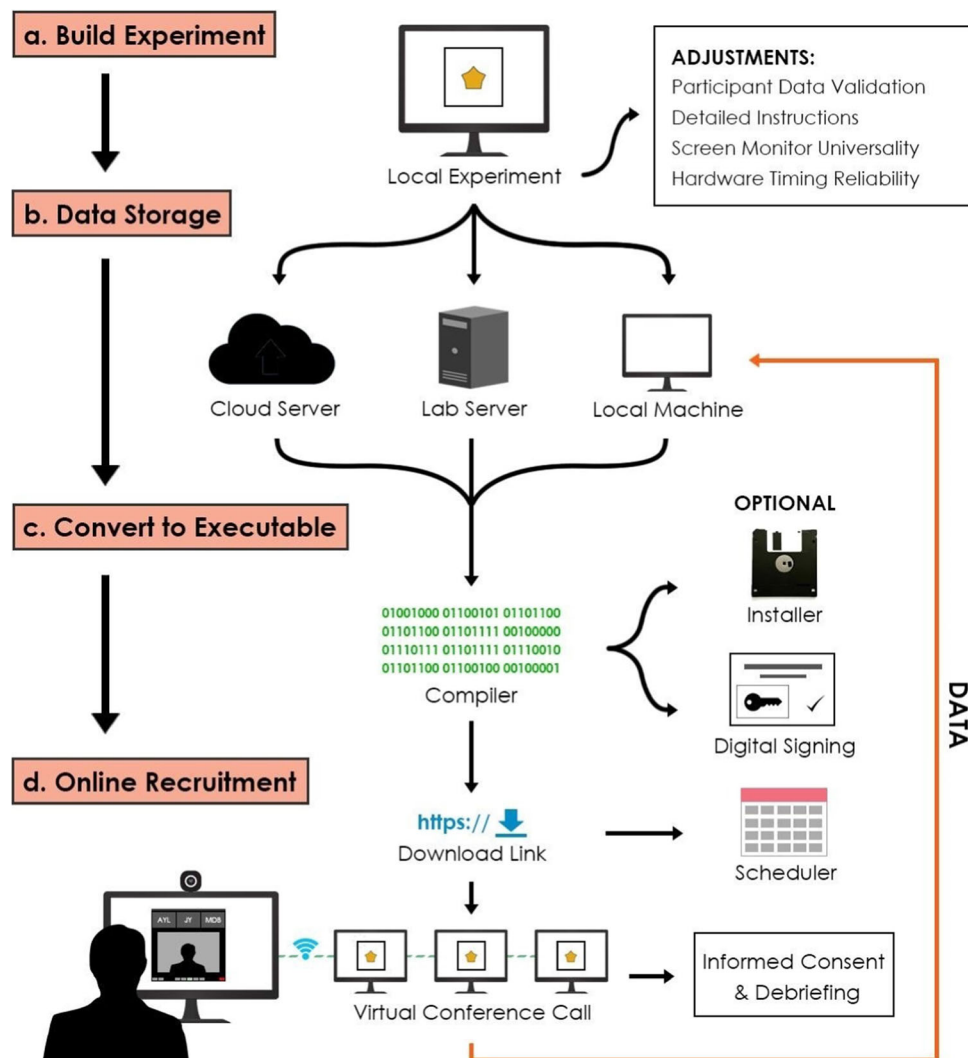
Fig. 2 Executable pipeline paired with online virtual conferencing. **a** The first step involves building a local experiment (e.g., using Python). Several adjustments to the experiment may be needed for testing participants online (see https://github.com/james-y-yuan/executable-pipeline for details). **b** To store the data, options include secure cloud services (e.g., Dropbox) or a lab server. Alternatively, executables can save data to the participant's local machine, and participants can manually upload the data to a secure cloud service. At this stage, researchers must carefully consider security and privacy concerns (see Discussion for further considerations). **c** To create the executable, open-source packages (e.g., PyInstaller) automatically compile the local experiment. To reduce concerns with installation and firewalls blocking the program, the experimenter can also include an installer as well as digitally sign the program. **d** Finally, the executable can be downloaded, installed, and run by participants. In our study, participants ran the task on their own machines during a synchronous virtual call to improve data quality at the beginning of our experiment, but this step could be completed asynchronously depending on the research question. (Color figure online)

many participants in groups within the concurrent testing condition (Fig. 3c). Having established the efficacy of our online approach, we then conducted mouse-tracking trajectory analyses to understand how participant reconstruction behavior may be influenced by memory load. Finally, we examined the influence of memory load on object-to-location binding errors during object reconstruction. For details about converting Python-based experiments into a downloadable executable, including a full video demo, see a tutorial and commented code on GitHub: https://github.com/james-y-yuan/executable-pipeline.

Participants

Sixty participants were recruited from the University of Toronto and from the community (M_{age} = 20.41 years, SD_{age} = 3.83 years, 31 females). Thirty participants (M_{age} = 21.97 years, SD_{age} = 4.90 years, 20 females) were tested individually (Fig. 3a), and 30 participants (M_{age} = 18.90 years, SD_{age} = 1.37 years, 12 females) were concurrently tested (Fig. 3b). Students received course credit, and community members were compensated with $10/hour CAD.
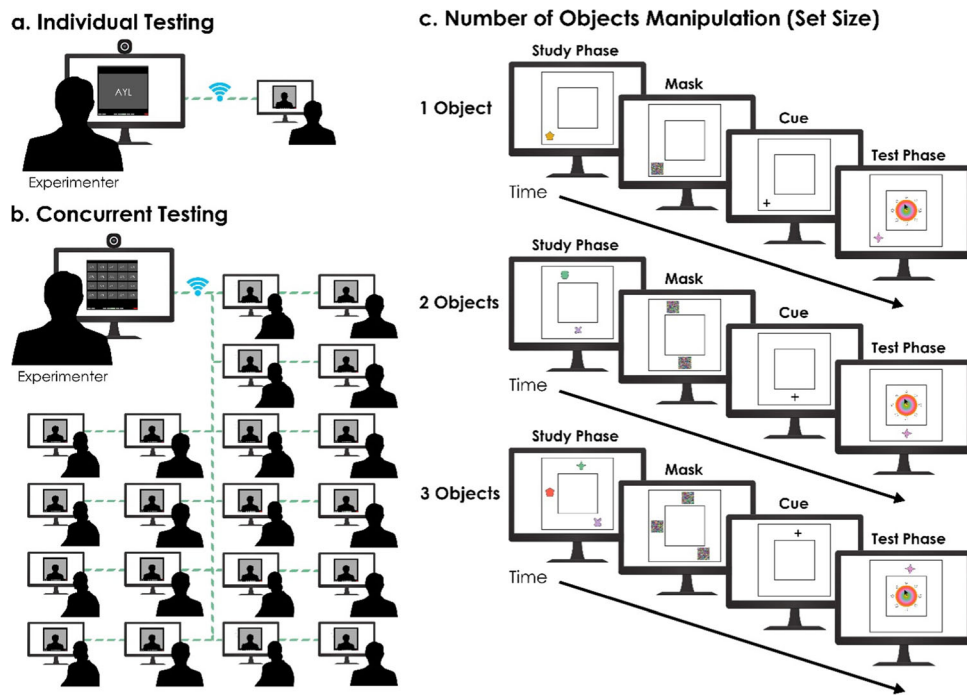
Fig. 3 Experimental procedure. Participants across both (**a**) individual testing and (**b**) concurrent testing studied to-be-remembered shape–color objects in (**c**) trials that varied in memory load (one object, two objects, or three objects). During the study phase (2,000 ms), participants studied either one, two, or three objects. The objects were randomly sampled from a minimum of 60 errors on shape and color space each, such that visual similarity was explicitly controlled. We compared individual testing (**a**), where a single participant completed the task while they were in a virtual conference room with the experimenter, to concurrent testing (**b**), where many participants completed the task while they were in the same virtual conference room with the experimenter. In this way, we examined whether we could scale data collection using concurrent testing in which many participants were present in the same virtual conference call (i.e., over Zoom), and whether a downloadable executable method of online testing could replicate previous memory findings from the lab (A. Y. Li et al., 2022). Furthermore, we could examine whether binding errors occur throughout object reconstruction using a novel online mouse-tracking task. (Color figure online)

Procedure

We incorporated mouse-tracking into the simultaneous reconstruction task, such that the position of the mouse cursor was recorded approximately once every 20 ms (for more details about the task design, see A. Y. Li et al., 2022; Fig. 1). Shapes were sampled from the Validated Circular Shape Space (A. Y. Li et al., 2020), and colors were derived from a circle defined on CIELAB *color space* (L = 70, a = 20, b = 38, radius = 60; Zhang & Luck, 2008). To ensure that there were no systematic mappings between particular features and locations on the task, we jittered shape mappings by participant and color mappings by trial, following previous work (A. Y. Li et al., 2022).

To manipulate memory load, participants were presented with twenty trials each of one, two, or three shape–color objects in a random order (Fig. 3c). Objects were displayed within a fixed 1,080 × 1,080-pixel square, coded using absolute coordinates so that visual images were never stretched or distorted across different monitor resolutions (see Fig. 3c). Object locations were sampled randomly so that no objects overlapped any other objects. The shapes and colors of the objects were sampled from VCS space and CIELAB space, respectively. When sampling from each feature space, values for a given trial were chosen from a set of six points spread equidistantly across the circle (i.e., 60 degrees apart). Thus, for every trial, all objects were at least 60 degrees different in shape and color from every other object, so that the visual similarity of objects was always tightly controlled. This sampling approach also ensured that the target object's shape and color features were random for each trial.

Each trial proceeded in the typical fashion of continuous reconstruction tasks (see Fig. 1a). After the initial study phase of 2,000 ms, there was an ISI of 300 ms. The masks then appeared at the location of the studied objects for 300 ms. After a retention interval of 1,000 ms, a cue appeared at the location of one of the studied objects (500 ms). The test phase display then appeared, with the mouse cursor positioned at the centre of the screen. Participants used their mouse to reconstruct the target shape–color object which was cued at an indicated location (Fig. 3c). At the onset of the initial mouse movement, the reconstructed object appeared at the cued location corresponding to the position of the mouse cursor on the test display. This portion of the task was untimed, and we recorded the mouse movement throughout the entire test phase until the participant made a response (Fig. 1b). Thus, participants complete each trial by matching their memory of

the target object with what was perceived on the display as the mouse cursor moved along shape and color space (for a video, see: https://osf.io/ycq5s). Upon completion of the task, which ranged from 30 to 60 minutes, the anonymized data was uploaded to our lab server (Fig. 2), and participants were debriefed verbally and compensated. To ensure that our task could be reasonably completed online within this time frame, each set size condition included 20 trials (see Discussion for more information about this methodological decision).

In the individual testing condition, we completed the task across 30 sessions of one-on-one virtual conference calls (Fig. 3a). Critically, we completed the task across six sessions of multiple groups of participants in the concurrent testing condition (Fig. 3b). We tested up to 18 participants in the same virtual conference, with the other sessions containing up to five participants. The concurrent testing condition was intended to mimic a typical online study, where an experimenter may wish to recruit many participants on a rolling basis.

Statistical analysis

Online task data reliability

To determine the data reliability of our online executable pipeline using the simultaneous reconstruction task (Fig. 1), we first examined whether we could replicate previously observed memory load effects from the lab (e.g., A. Y Li et al., 2022). We then tested the efficiency of our virtual conference room approach by comparing memory performance between the individual and concurrent testing conditions (Fig. 3).

Memory performance was quantified as *error*, defined as the absolute angular distance between the reconstructed feature and target feature on the individual circular shape and color spaces (see Fig. 1b). More specifically, we defined *fine-grained feature responses* as those for which shape or color error were less than or equal to 15 degrees. Furthermore, we defined *fine-grained object responses* as those for which both shape *and* color error were less than or equal to 15 degrees (see A.Y. Li et al., 2022). We predicted that increases in memory load should lead to decreases in fine-grained responses, in line with previous findings in the literature (A. Y. Li et al., 2022; Ma et al., 2014; Sone et al., 2021).

Separate random-intercept linear mixed models were used to predict fine-grained shape, color, and object responses from memory load (see Magezi, 2015, for how within-subject psychology experiments can be analyzed using linear mixed models). For each model, fine-grained shape, color, or object responses were used as the dependent variable, and memory load (Set Size 1, 2, or 3) was defined as the fixed factor. In a follow-up analysis, to directly compare between individual and concurrent testing conditions (Fig. 3), we additionally included testing condition as a fixed effects factor. In all

analyses, individual participants were included as a clustering variable to account for the within-subject design. All linear mixed model analyses were conducted using *jamovi* (The Jamovi Project, 2021).

Mouse-tracking trajectory analysis

Once the reliability of our online approach had been established, we then conducted a series of mouse-tracking trajectory analyses. We first calculated descriptive results, including the onset time of the mouse movement as well as the mean and variability of the trial duration across mouse trajectories. We next analyzed the continuous trajectory, including only the responses when the mouse cursor began moving and not when the cursor was displayed at the centre of the screen at the start of the test phase.

Because we recorded the position of the mouse cursor approximately once every 20 ms, and because the duration of trials was not fixed, the number of data points across trials will vary based on their duration (e.g., a 1,000 ms trial would give us 50 data points, whereas a 2,000 ms trial would give us 100 data points). Thus, even two paths that follow an identical trajectory over shape–color space will differ in the number of data points collected if their durations are not the same (see Fig. 5a for a visual depiction). For this reason, we linearly interpolated each trajectory to the same path length to normalize the duration of all mouse paths (Fig. 5b). This approach allowed us to project the trajectories across all trials of the experiment onto the same axis. We then converted the shape–color object corresponding to the position of a given mouse cursor at each time point into error (i.e., how far that position was from the target position of the shape–color object; Fig. 1b). By depicting the normalized trajectories in terms of error from the target, we could graphically visualize participant object reconstruction behavior over relative positions along the trajectory (Fig. 5). Put simply, we visualized each point along the trial in a manner akin to how the final response is typically depicted (Fig. 4).

In the final trajectory analysis, we analyzed how the normalized mouse paths changed across relative positions along the trajectory as a function of memory load (Fig. 6). We visualized the mouse trajectories for each participant in terms of shape errors and color errors separately, extending our previous visualization which projected errors along shape–color space together onto the same axis (Fig. 5). The dependent variable in our linear mixed models was the shape error and color error of each data point along the normalized mouse trajectory. Fixed factors were memory load (one objects, two objects, three objects) and relative position in the trial (start, 25%, 50%, 75%, and final response). This approach allowed us to statistically assess the magnitude of the error across memory load conditions and the defined positions along the trajectory. As
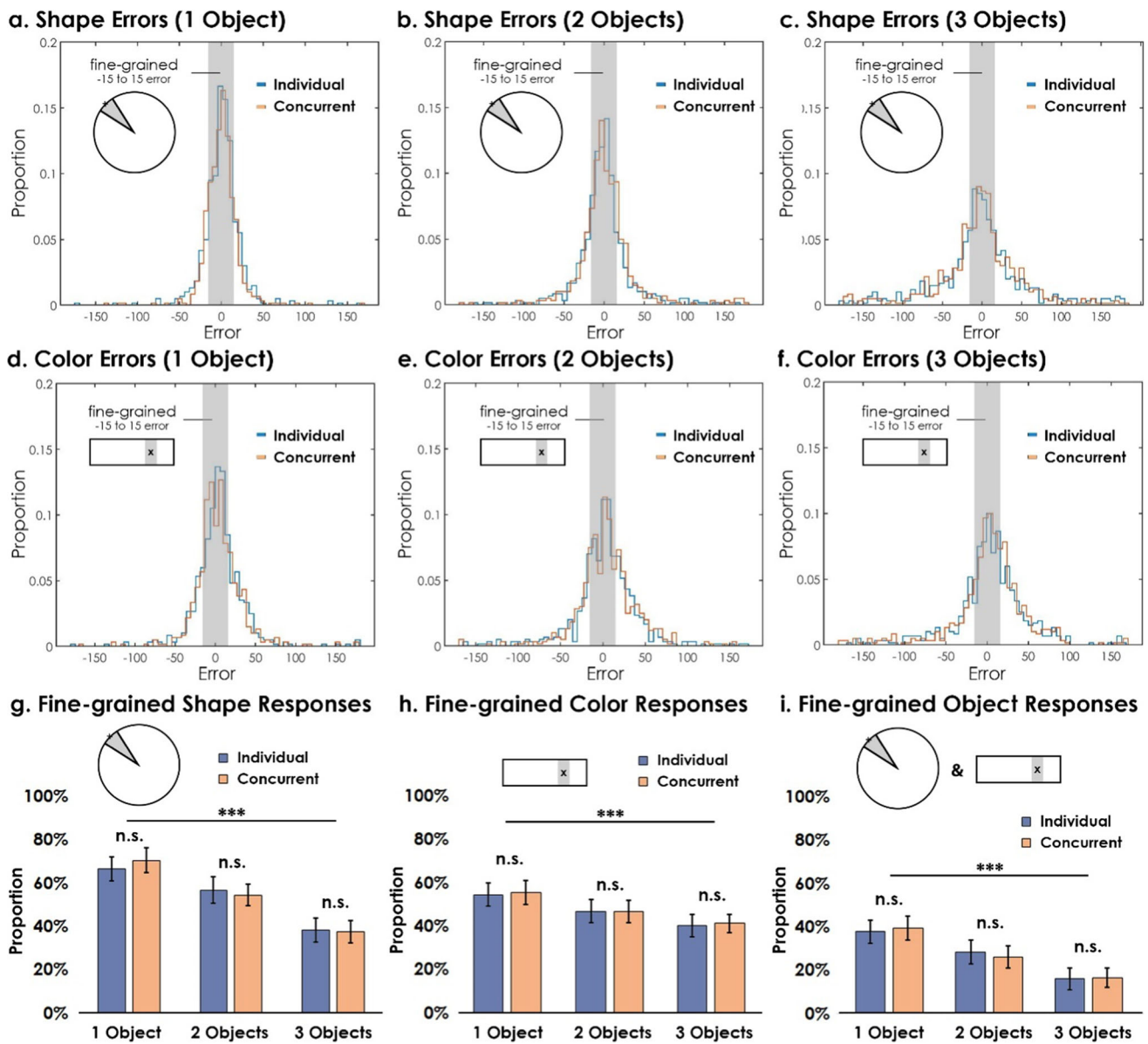
Fig. 4 Results from individual and concurrent testing manipulations. **a–f** Qualitatively, there was little difference between the overall response error distributions between conditions for the shape and color features across each method of testing. **g** We replicated a previously described memory effect, such that fine-grained shape (**g**), color (**h**), and object responses (**i**) decreased as the memory load increased (all $ps < .001$). Critically, there was no difference in any comparison between individual testing compared with concurrent testing, with all fine-grained responses

replicated within 4% across conditions. These results demonstrate the data reliability of online executables when paired with a synchronous virtual conference call. Furthermore, these results suggest that testing many participants in the same virtual call during concurrent testing can be a scalable method that increases testing efficiency. Error bars refer to 95% confidence interval for the mean. *** refers to $p < .001$ and *n.s.* refers to not significant. (Color figure online)

clustering variables, we included information about relative position, individual trials, and participants.

Quantifying object-to-location binding errors

Next, we examined whether memory load influences object-to-location binding errors during object reconstruction; see Fig. 7. For each trial, we identified the number of times a participant's mouse hovered over nontarget, uncued

objects presented during the study phase (i.e., a nontarget lure object; Fig. 7a). More specifically, any period in time, of any duration, in which the mouse cursor was within 15 degrees of both shape and color error from a nontarget object was defined as a "nontarget mouse hover" (see example: https://osf.io/7tfn6/). We report this analysis only for trials with one nontarget object (Set Size 2) and two nontarget objects (Set Size 3), as only these trials presented participants with nontarget objects. The benefit
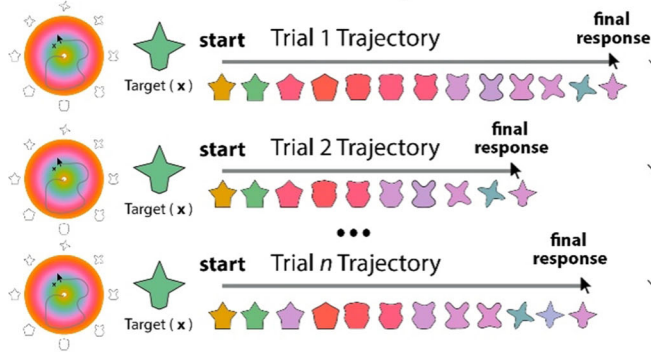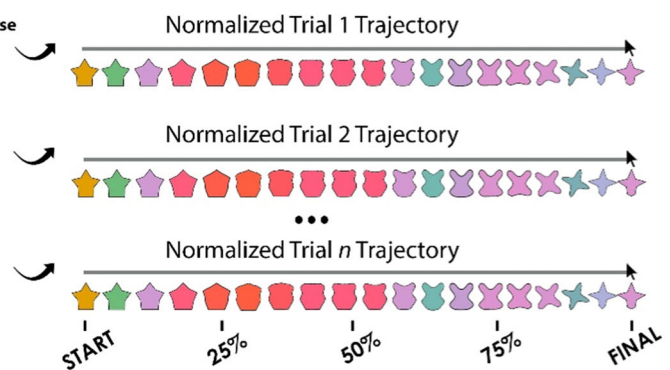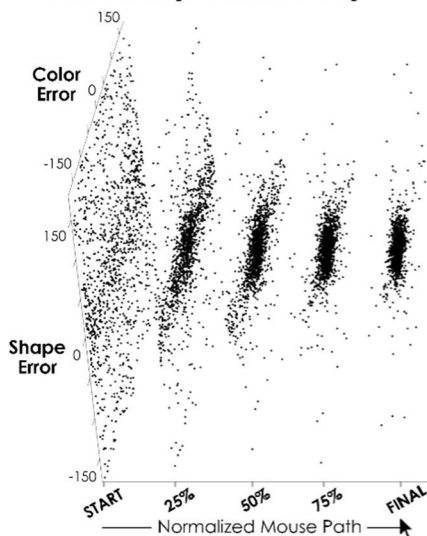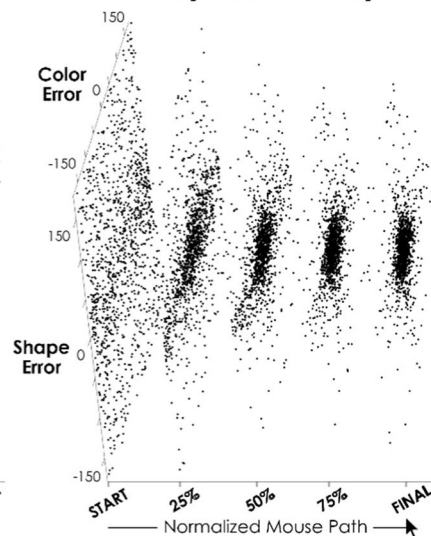
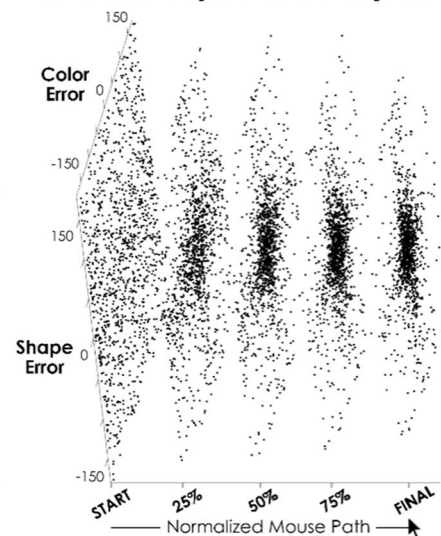**Fig. 5** Mouse-tracking trajectory analysis. **a** Example trials with identical trajectories but different duration. Because data points are collected every 20 ms, even identical trajectories will produce a different number of data points if the mouse travels at different speeds. **b** For this reason, we used linear interpolation to normalize all trials of the experiment so that they could be visualized together onto the same axis. We then converted the shape–color object corresponding to the current position of the mouse cursor at each time point as *error* from the target shape–color object (Fig. 1b). **c–e** This approach produces multivariate error distributions for each set size across relative positions along the trajectory (start, 25%, 50%, 75%, final). As displayed above, errors along shape–color space gradually cluster near zero error as the mouse trajectory reaches the final response. Mouse trajectories are visualized as shape error (*y*-axis) and color error (*z*-axis) in terms of the relative position along the trajectory (*x*-axis). (Color figure online)

of this approach is that we can examine the trajectory leading up to the final response at test, akin to the analysis of eye-tracking experiments that quantify the interference between target and nontarget object lures (Barense et al., 2012; Erez et al., 2013; Yeung et al., 2013; Yeung et al., 2017; Yeung et al., 2019). Thus, mouse tracking provides a potentially sensitive marker of object-to-location binding errors that occur during object reconstruction.

We used the same linear mixed models described previously, except we predicted nontarget hovers from memory load while including potentially confounding factors in the model. We included trial duration as a fixed effect, as mouse paths that span longer durations tend to move over broader areas of the available shape–color space, inflating the chance of

random hovers over nontargets. We also included the overall variability of the mouse path trajectory over shape and color spaces as fixed effects, which would similarly cause systematic increases in hovers due to mouse paths spanning a broader area.

Ruling out alternative explanations

To ensure that our measure of nontarget mouse hovers reflects actual object-to-location binding errors that increase with memory load rather than random hovers driven by chance, we accounted for several alternative explanations. First, our operationalization of mouse hovers is systematically influenced by the number of objects in a trial. This is because

mouse hovers will be more likely to occur by chance on trials with two nontargets (Set Size 3) compared with one nontarget (Set Size 2), given that twice as much area counts as a hover when two nontargets are present (see Fig. 7a). Thus, if more hovers are observed in Set Size 3 than in Set Size 2, this could be attributed merely to more mouse movements being captured as nontarget hovers by chance alone. For this reason, we conducted a control analysis that equated the area that could be counted as a nontarget hover across set sizes. Using the same sampling parameters as the actual nontargets (see Procedure), we generated an "invisible" nontarget in the Set Size 2 trials and then counted the number of mouse hovers over both the *actual* nontarget and the *invisible* nontarget. By doing so, the area counted as a hover in Set Size 2 is made equal to the area in Set Size 3.

Next, we accounted for the theoretical possibility that an increase in mouse hovers over nontarget objects might be driven by random guessing behavior. Whereas the previous analysis controlled for the random hovers driven by chance from our *analytical approach*, the present analysis controlled for the random hovers driven by chance from *participant mouse movements*. That is, perhaps more hovers could be observed at higher memory loads not because of binding errors, but because participants tend more often to forget the target object and resort to a guess (e.g., Zhang & Luck, 2008, 2009). This possibility is usually addressed by fitting a uniform distribution from a mixture model, derived from the final responses at test across all trials of an experiment. Here, we addressed the possibility of random guessing behavior trial-by-trial during object reconstruction. If participants select a random shape–color object on the stimulus space (Fig. 7a), we expect to find more hovers at any given point on the shape–color space, not only over nontargets. Thus, we counted the number of hovers over the points directly *opposite* to nontarget objects on shape–color space each trial, which should increase with memory load only if random guessing is involved. On the contrary, however, we expected that increases in nontarget mouse hovering could not be explained by random guesses at higher memory loads, and so we predict that hovers over *nontargets* should be unrelated to hovers over *a point on the opposite side* of shape–color space. Thus, in a subsequent linear mixed model control analysis, we included the hovers over the opposite side of shape–color space as a fixed effect to control for the possibility of random guessing behavior.

Finally, we examined whether object-to-location binding errors occurred throughout the entire period of object reconstruction, providing a rich characterization of participant behavior leading up to the final response at test. We compared the frequency of hovers per trial in the first 75% of the trial's mouse path (i.e., object reconstruction) to the final 25% of the trial's mouse path (i.e., the final response at test). As longer mouse path lengths provide more opportunities to hover over

nontarget objects, we normalized the hover frequencies to ensure the results were in proportion to the entire mouse path. Specifically, we multiplied the hover frequency over the first 75% of the mouse path by 4/3, and the frequency over the final 25% of the mouse path by 4, so that all hover frequencies we report in the results are in proportion to 100% of the mouse path.

Results

For anonymized data, see https://osf.io/a4vsb/, and a tutorial for creating downloadable executables is available at https://github.com/james-y-yuan/executable-pipeline. See the raw error distributions for shape and color features from the final response at test in Fig. 4a–f (i.e., the raw error distributions akin to a typical continuous reconstruction task; Ma et al., 2014), and the error distributions over the entire mouse trajectory in Fig. 5 and Fig. 6.

Online task data reliability

Using online executables paired with virtual conference rooms, we first aimed to replicate a previous effect from in-person testing using the simultaneous reconstruction task—namely, that fine-grained responses decrease as memory load increases (A.Y. Li et al., 2022; Ma et al., 2014; Sone et al., 2021). To assess the testing efficiency of our approach, we also compared individual testing, where participants completed the task one-by-one with the experimenter (Fig. 3a), to concurrent testing, where many participants completed the task in the same virtual conference call (Fig. 3b).

We used separate linear mixed models to predict fine-grained shape responses, fine-grained color responses, and fine-grained object responses from set size in the individual testing condition. Fine-grained feature responses were calculated for the shape feature and color feature independently, and fine-grained object responses were calculated for both shape and color features together. We found that participants were less able to reconstruct the exact target shape (Fig. 4g), target feature (Fig. 4h), or target shape–color object (Fig. 4i) as the number of studied objects increased. For the individual-testing condition, there was a significant main effect of memory load on the proportion of fine-grained shape responses, $F(2, 58) = 11.78$, $p < .001$, $\eta_p^2 = 0.29$; fine-grained color responses, $F(2, 58) = 48.48$, $p < .001$, $\eta_p^2 = 0.63$; and fine-grained object responses, $F(2, 58) = 28.99$, $p < .001$, $\eta_p^2 = 0.50$ (Fig. 4g). These results suggest that testing participants online with executables can produce reliable data, as we replicated previous findings from the lab (see A.Y. Li et al., 2022).

Revealing the testing efficiency of concurrent virtual conference calls, we were able to replicate the findings from the
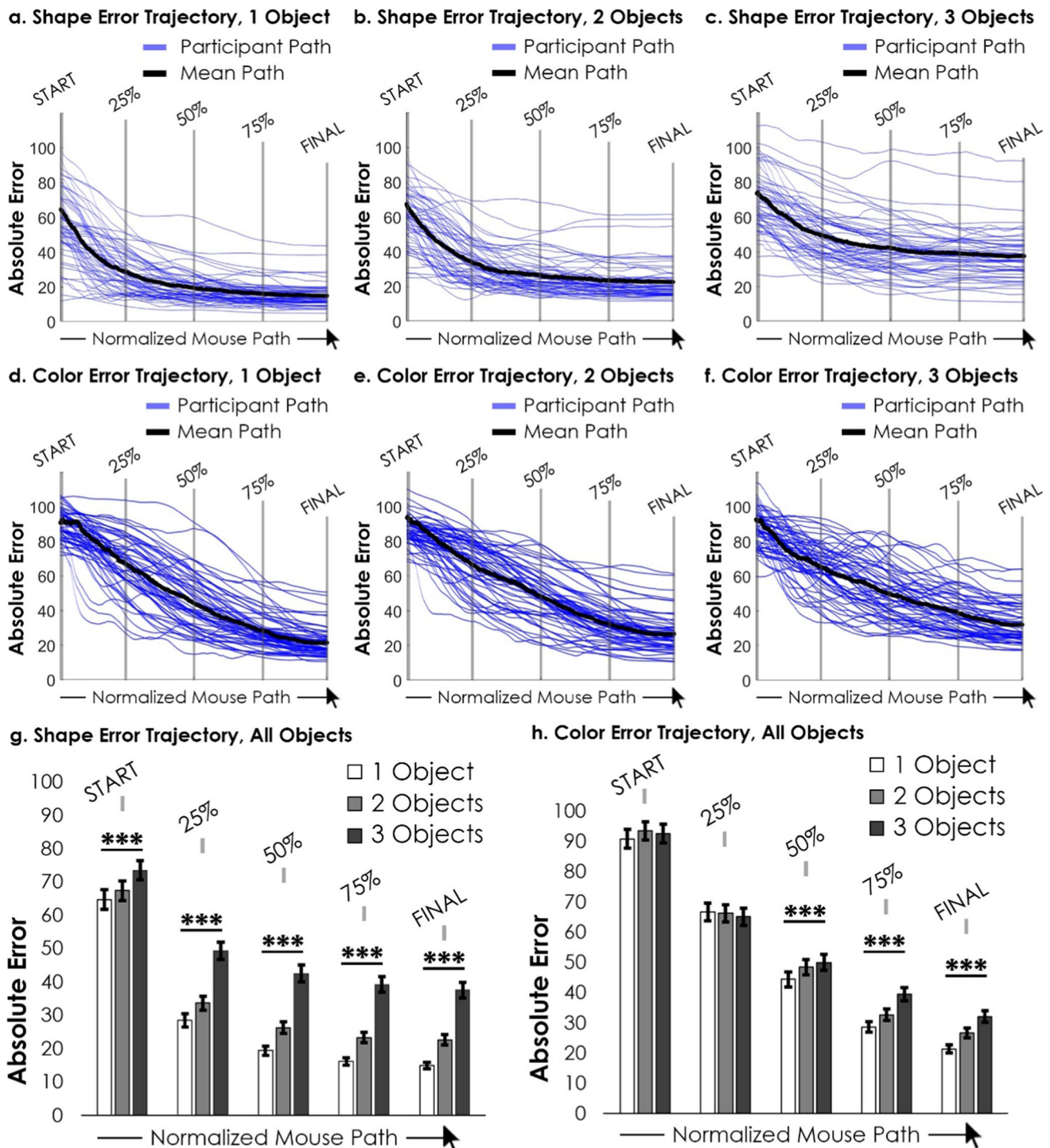
Fig. 6 Shape and color error trajectories are shown separately as a function of memory load. During object reconstruction, participants initially moved rapidly along shape (**a, b, c, g**) and color space (**d, e, f, h**) before slowing down, an effect that interacted with memory load and the relative positions along the trajectory (**g, h**). Critically, because distance on circular space is a proxy for visual similarity, we know that participants moved more rapidly over dissimilar shapes and colors (i.e., higher error) compared with similar shapes and colors (i.e., lower error).

Indeed, the slope of errors from the first half of the trajectory differed from the final half of the trajectory and was influence by memory load. These results provide evidence that decision uncertainty was partially driven by interference between working memory (i.e., the target object in memory) and the similarity of the currently perceived object (i.e., the reconstructed object corresponding to the current position of the mouse cursor). Error bars refer to 95% confidence interval for the mean. *** refers to $p < .001$. (Color figure online)

individual testing condition within 4% of values for all fine-grained responses (see Fig. 4). For the concurrent testing condition, we observed a significant main effect of number of objects on the proportion of fine-grained shape responses, $F(2, 58) = 67.60$, $p < .001$, $\eta_p^2 = 0.70$; fine-grained color responses, $F(2, 58) = 13.20$, $p < .001$, $\eta_p^2 = 0.31$; and fine-grained object responses (Fig. 4h), $F(2, 58) = 40.97$, $p < .001$, $\eta_p^2 = 0.59$. Critically, when we included the individual and concurrent testing conditions within the linear mixed model as a fixed factor, we found no difference between the individual and concurrent testing conditions for fine-grained shape memory, $F(1, 58) = 0.20$, $p = .66$, fine-grained color memory, $F(1, 58) = 0.019$, $p = .89$, or fine-grained object memory, $F(1, 58) = 0.010$, $p = .75$ (Fig. 4g, h, i).

These results not only suggest that data from online testing can be reliable when paired with synchronous virtual conference calls, but also provide evidence that concurrent testing can be a scalable way to increase experimental efficiency while maintaining high data quality during online testing. See further validation measures when participants are tested with downloadable executables on GitHub: https://github.com/james-y-yuan/executable-pipeline.

Mouse-tracking trajectory analysis

One participant was excluded as their mouse paths were not recorded due to technical error. Notably, we found no significant difference in trial duration across memory load conditions (one object: *Mean* = 6.42 seconds, *SD* = 3.67 seconds; two objects: *Mean* = 6.08 seconds, *SD* = 3.37 seconds; three objects: *Mean* = 6.12 seconds, *SD* = 3.73 seconds), $F(1, 2359) = 3.07$, $p = .08$. There was also no significant difference in the onset time of the mouse movement across memory load conditions (one object: *Mean* = 0.59 seconds, *SD* = 0.28 seconds; two objects: *Mean* = 0.57 seconds, *SD* = 0.32 seconds; three objects: *Mean* = 0.62 seconds, *SD* = 0.43 seconds), $F(1, 2359) = 1.65$, $p = .20$.

Next, we looked at the entire mouse-path trajectory during object reconstruction (Fig. 5a), enabling us to study participant behavior leading up to the final response at millisecond temporal resolution. To visualize mouse paths across memory load conditions onto the same axis, we linearly interpolated the trial-by-trial trajectories (Fig. 5a) to be of the same relative path length (Fig. 5b). We then plotted absolute shape error (*y*-axis) and color error (*z*-axis) from the target by relative position along the trajectory (*x*-axis), allowing us to visualize participant mouse paths in terms of continuous error to the target (Fig. 5c–e). This approach extends the analysis of error distributions from only the final response at test (e.g., Ma et al., 2014; Fig. 4a–f) into the entire continuous mouse-path trajectory (Fig. 5c–e). In other words, we used mouse-tracking to convert a dependent variable with a single observation each trial (i.e., the final response

on a continuous reconstruction task) into time-series data with hundreds of observations each trial (i.e., the entire continuous mouse path trajectory).

We next displayed the mouse trajectories by shape errors and color errors individually for each participant (Fig. 6). Although the average trial duration and the onset time did not differ across memory load conditions (see previous analysis), the magnitude of errors robustly differed across memory load conditions (Set Size: one object, two objects, and three objects) and across relative positions along the trajectory (i.e., absolute error at: onset, 25%, 50%, 75%, final response). We found a main effect of memory load, shape: $F(1, 16511) = 926.42$, $p < .001$; $\eta2p = 0.05$; color: $F(1, 16511) = 49.93$, $p < .001$; $\eta_p^2 = 0.0042$, a main effect of position along the trajectory, shape: $F(1, 16511) = 937.68$, $p < .001$; $\eta_p^2 = 0.44$; color: $F(1, 16511) = 1432.44$, $p < .001$; $\eta_p^2 = 0.55$, and an interaction between memory load and relative position along the trajectory, shape: $F(1,16511) = 18.12$, $p < .001$; $\eta_p^2 = 0.0061$; color: $F(1, 16511) = 10.21$, $p < .001$; $\eta_p^2 = 0.0035$. In other words, we found that (i) errors for shape and color were higher when participants studied more objects, (ii) errors were lower as participants neared their final response during object reconstruction, and (iii) the slope of errors across relative positions along the trajectory differed by memory load (i.e., interaction between memory load and position along the trajectory; Fig. 6).

These results suggest that participants initially moved rapidly over shape–color space before slowing down as they neared the target, and this effect was modulated by memory load. Because distance on circular space is a proxy for visual similarity (e.g., validated in previous experiments; for shape, see A. Y. Li et al., 2020; for color, see Schurgin et al., 2020), these results reveal that participants moved rapidly over dissimilar objects (i.e., high error) before slowing down over similar objects (i.e., low error) as they neared their final response. That is, the fact that mouse movement slowed as the objects became more similar to the target during reconstruction suggest interference between the contents of working memory (i.e., the target object held in mind) with the currently perceived object (i.e., the object on the display corresponding to the position of the mouse cursor; Fig. 1a). By analogy to eye tracking, our results would suggest that participants move their eyes rapidly over dissimilar objects relative to the target before slowing down in their gaze behavior when objects are similar to the target.

Quantifying object-to-location binding errors using mouse tracking

The previous analyses examined how memory load influenced object reconstruction for the *target* shape–color object, finding interference between the contents of working memory (i.e., the target object) with the similarity of the viewed object

over time (i.e., the object corresponding to the position of the mouse cursor). In our next analysis, we looked at how memory load might influence object-to-location binding errors, providing evidence of between-item interference stemming from previously studied but currently irrelevant nontarget object lures. Inspired by previous eye-tracking methodologies (Erez et al., 2013; Ryan et al., 2000; Yeung et al., 2013; Yeung et al., 2017; Yeung et al., 2019), we operationalized object-to-location binding errors as the number of mouse hovers over nontarget objects from the study phase (Fig. 7a).

We used a linear mixed model to predict hovers over nontarget objects from set size, controlling for the potential confounding factors of trial duration and variability of the shape and color paths. We found that memory load significantly predicted the number of hovers over nontarget objects, $F(1, 113) = 121.88$, $p < .001$ (Fig. 7b), such that participants were more likely to mouse over nontarget objects during the condition with two nontarget objects at Set Size 3 ($M = 0.27$ hovers per trial, $SD = 0.18$) compared with the condition with one nontarget object at Set Size 2 ($M = 0.10$ hovers per trial,

a. Quantifying object-to-location binding errors as non-target mouse hovers
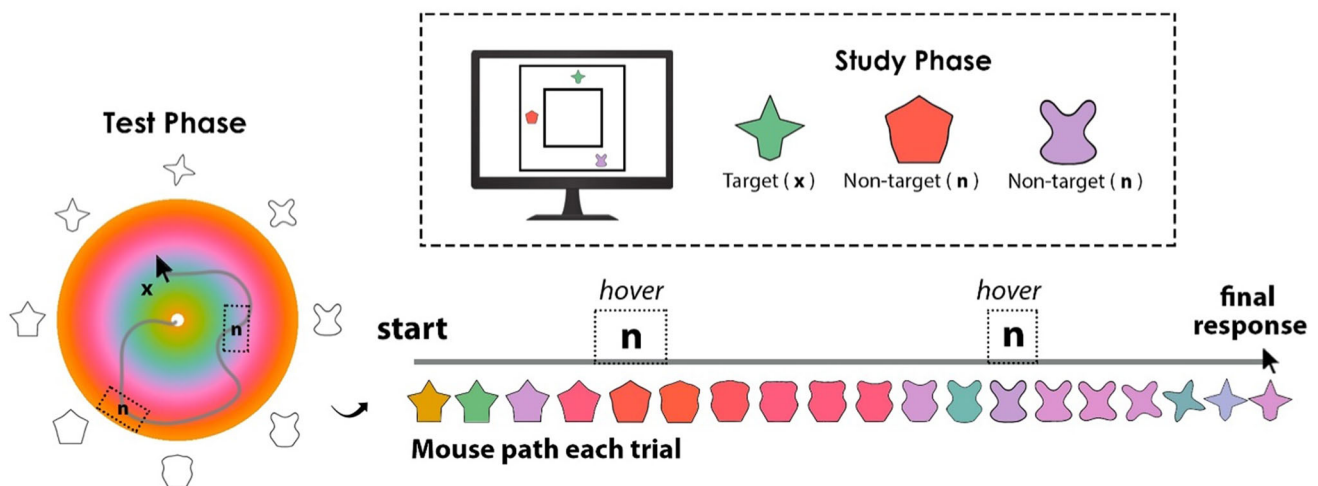
b. Non-target mouse hovers during object reconstruction increase with memory load
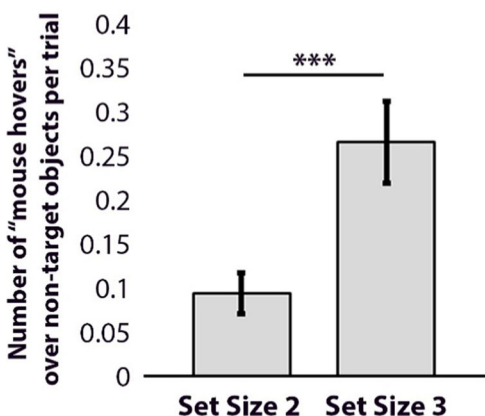
c. Non-target mouse hovers occur more often during object reconstruction than the final response, normalized by path length
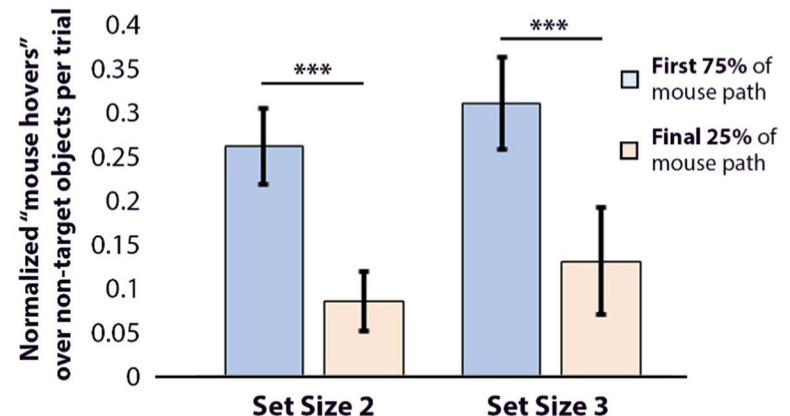
Fig. 7 Mouse-tracking analysis of object-to-location binding errors (i.e., mouse hovers over nontarget object lures). **a** From the mouse path on each trial, we counted the number of "mouse hovers" over nontarget objects as participants reconstructed the target object's shape and color. A hover was defined as a mouse movement, of any length of time, over 15 errors around a nontarget object's shape and color. **b** In the initial analysis, we found that participants were nearly 3× more likely to mouse hover over nontarget object features at the highest memory load. Subsequent control analyses ensured that this effect was not driven by random chance stemming from our operationalization of nontarget mouse hovers nor by random guessing behavior (see Results). **c** When we

separately analyzed the first 75% of the mouse path (i.e., object reconstruction behavior) from the final 25% of the mouse path (i.e., the final response at test), we found that participants were 2.64 times more likely to make nontarget mouse hovers during object reconstruction. This comparison was normalized by path length, so that the hover frequencies in (**c**) account for path length differences across object reconstruction and the final response. In summary, these results provide evidence of between-item interference stemming from previously studied but currently irrelevant lures during object reconstruction. Error bars refer to 95% confidence interval for the mean. *** refers to $p < 0.001$. (Color figure online)

$SD = 0.093$). Participants were almost three times more likely to mouse over nontarget object features at the highest memory load, suggestive of object-to-location binding errors that occurred throughout object reconstruction (i.e., the influence of previously studied but currently irrelevant objects). The variability of the mouse movement in shape space significantly predicted the number of hovers over nontargets, $F(1, 113) = 4.80$, $p = .031$, but not the variability of the mouse movement in color space, $F(1, 113) = 0.62$, $p = .43$, or the length of time it takes to complete a trial, $F(1, 113) = 0.25$, $p = .62$.

Ruling out alternative explanations

In the subsequent analyses, we addressed several alternative theoretical possibilities (for rationale, see Statistical Analysis). To ensure that our results could not be driven by random chance driven by our operationalization of nontarget mouse hovers, we generated an "invisible" object in the trials with one nontarget object (Set Size 2) to equate area to the trials with two nontarget objects (Set Size 3). In addition, to account for the possibility that more hovers at higher set size might be explained by random guessing behavior, we also included the number of hovers over the opposite side of shape–color space from the nontargets into a linear mixed model. That is, if random guessing increased with set size and entirely accounted for our results in Fig. 7b, then these opposite-side hovers should also increase in tandem with nontarget hovers. Even after these stringent controls, participants were still 22% more likely to mouse over nontargets at the highest memory load during object reconstruction, $F(112) = 7.01$, $p = .009$. That is, these control analyses assume that participant mouse behavior was driven by random chance, yet we still observed evidence that memory load influenced mouse hovers over nontarget object features. Thus, 22% reflects the lowest possible difference in object-to-location binding errors between Set Size 2 and Set Size 3 trials after assuming that mouse hover behavior stem from random chance. By contrast, trial duration, $F(1, 112) = 2.48$, $p = .12$, standard deviation of shape path, $F(1, 112) = 0.93$, $p = .34$, and standard deviation of color path, $F(1, 112) = 0.73$, $p = .40$, did not significantly predict the number of hovers over nontarget object features. Critically, hovers over the *opposite side* of shape–color space did not increase in tandem with *nontarget object hovers*, $F(1, 112) = 0.16$, $p = .69$. Thus, our results are not explained by random chance stemming from our operationalization of nontarget mouse hovers. Nor are they explained by random guessing behavior from our participants, as then the number of hovers over locations opposite to nontargets should have increased in tandem with hovers over nontarget objects.

Finally, we calculated the average number of hovers observed per trial using only the first 75% of each trial (i.e., during object reconstruction) compared with only the final 25% of each trial (i.e., the final response at test). After normalizing the responses to the length of the mouse path each trial, we found that the average number of hovers over nontargets in the first 75% of trials ($M = 0.29$ hovers per trial, $SD = 0.19$) exceeded the average number of hovers over nontargets in the final 25% of the path by nearly a factor of 3 ($M = 0.11$ hovers per trial, $SD = 0.20$), $t(234) = 7.25$, $p < .001$; Fig. 7b. Participants were 2.64 times more likely to make mouse hovers over nontarget object features during the first 75% of the mouse path compared with the final 25% after correcting for path length, revealing that mouse reconstruction behavior can provide a sensitive marker of between-item interference from previously studied but currently irrelevant object lures.

Overall, our mouse trajectory analyses reveal that object reconstruction behavior provides rich information about the internal representation leading up to the final response on a continuous memory task. We observed interference between the contents of working memory with the currently perceived object (Figs. 5, 6), as well as from previously studied but currently irrelevant object lures (i.e., object-to-location binding errors; Fig. 7).

Discussion

By implementing mouse-tracking during an online executable version of a simultaneous reconstruction task, we show that memory load influences continuous mouse trajectories during object reconstruction. We observed interference between the contents of working memory and perception, such that participant mouse movement was more rapid when the currently viewed objects were dissimilar (i.e., higher error) before slowing down when the currently viewed objects were similar (i.e., low error) to the target object (Figs. 5, 6). Furthermore, we observed robust interference between previously studied target and lure objects during reconstruction (i.e., object-to-location binding errors), such that participants were nearly 3 times more likely to "mouse over" lure object features at the highest memory load (Fig. 7b). This between-item interference from object lures occurred primarily during the earlier reconstruction phase rather than the period leading up to the final response at test, suggesting that mouse-tracking can provide a sensitive index of object-to-location binding errors. We complement these trajectory analyses with data reliability and testing efficiency metrics of our online mouse-tracking task using an executable pipeline, finding that testing many participants in the same virtual conference room (i.e., over Zoom) can be an efficient method to ensure high data quality (Figs. 2, 3, 4). Taken together, these results show that tracking the dynamic mouse movement on a continuous memory task during online testing can provide a rich characterization of participant behavior leading up to the final response at test.

Mouse-tracking as participants reconstructed shape–color objects over perceptually uniform space (Fig. 1b) revealed two critical insights. First, mouse-tracking converts tasks which would measure only a single observation each trial (e.g., yes/no on a discrete task or the final error response on a continuous reconstruction task) into tasks measuring hundreds of observations each trial at millisecond temporal resolution (see Hao et al., 2021; Park & Zhang, 2022, for previous mouse-tracking approaches applied to working memory). Even when the number of trials in an experiment may be limited due to time constraints inherent with online testing, our mouse-tracking analyses was well powered because information from the entire trajectory could be incorporated into the analyses (Figs. 5, 6). However, while 20 trials per set size was sufficient for our specific purposes, future work will need to consider whether other research questions will require more trials. Approaches involving model-fitting (e.g., for models applied to working memory, see Bays, 2016; Oberauer, 2021; Rademaker et al., 2018; Schurgin et al., 2020; Sutterer & Awh, 2016; Zhang & Luck, 2008; also see evidence accumulation models: Bornstein et al., 2017; Evans & Wagenmakers, 2020; Fradkin & Eldar, 2022; Krueger et al., 2017; Shenhav et al., 2018) or individual differences (e.g., Baker et al., 2020; Xu et al., 2017) will have different considerations. Due to our limited trial numbers, it is also difficult to directly compare our trajectory measures (i.e., object-to-location binding errors indexed as nontarget object hovers) to more traditional model-based approaches based on the final response at test (i.e., feature-to-location binding errors indexed as "swaps" on a mixture model; see Bays, 2016; Bays et al., 2009). Nevertheless, we believe that our operationalization of binding errors is complementary to how binding errors are typically defined from the final response at test.

Second, in addition to measuring well-established aspects of motor movement (Woodworth, 1899), mouse-tracking provides a powerful tool to measure decision uncertainty leading up to the final response on a cognitive task. On our task, participants needed to match what they were representing in mind with what they were viewing on the display during object reconstruction (Fig. 1; for a video example, see https://osf.io/ycq5s). We observed that participants rapidly moved towards the target in the first half of the trial, before slowing down as they neared the target in the final half of the trial and approached the decision (Fig. 6). We suspect that a major contributor to this slowing behavior was due to the increasing interference as participants neared the end of the trial. As distance on shape and color space is a proxy for visual similarity (for shape, see A. Y. Li et al., 2020; for color, see Schurgin et al., 2020), the similarity between the target (held in working memory) and the reconstructed object (perceived on the screen) increased parametrically as the participant neared the target during their reconstruction. The slowing effects suggest interference between the target object from

working memory with the similarity of the currently perceived object corresponding to the mouse cursor. Traditionally, interference is considered as arising largely from lures and targets presented at study, yet we observed interference between the contents of working memory and information presented during the response phase of the task itself (Figs. 5, 6). These findings extend previous mouse-tracking work finding between-item interference for colors (Hao et al., 2021; Park & Zhang, 2022) to more complex shape–color object representations, as well as show that interference may arise from the similarity match between what is being held in memory with what is shown on the screen.

Although we suggest mouse-tracking can measure mnemonic-perceptual interference, there are other important sources of variance that should also be considered. For example, motor movements have long been known to have an initial accelerating and later decelerating component (Woodworth, 1899). These "coarse-to-fine" adjustments are typical for motor movement trajectories (e.g., Marteniuk et al., 1987), and are likely present in mouse-tracking tasks as well. However, these motor movements cannot be the only driver of the slowing we observed as the participant reached their final response (Figs. 5, 6). There was a robust difference between trajectories across memory load conditions, even though the overall time it took to complete these trials did not differ. That is, the motor requirements across conditions in the test phase were identical (i.e., participants reconstructed a single shape–color object; Fig. 3c). What distinguished the conditions was the amount of material presented at study, indicating that it was differences in the contents of *memory* rather than motor requirements that guided trajectory differences. For this reason, we suggest that interference between the contents of memory and what is perceived contributes to mouse trajectory behavior above and beyond motor movements alone. That is not to say that interference is the only factor that drives trajectory behavior—future research could more directly decouple the multiple sources of variance that contribute to the final response at test, potentially using different feature spaces that manipulate the similarity match between what is held in memory with what is perceived. Such an approach could also record response trajectories that do not involve mouse tracking, isolating the contribution of motor movements (e.g., participants could "hold down" arrow keys or move a joystick to reconstruct objects across the dimensions of perceptual space). Finally, this line of research could examine the interaction between trajectory speed and target–lure similarity (Figs. 5, 6) in other domains involving trajectories such as gaze behavior or spatial navigation. These future investigations may offer exciting avenues to further study the interaction between memory and perception.

Notably, our continuous measure of mouse tracking may be analogous to previously described measures of eye tracking (e.g., Barense et al., 2012; Erez et al., 2013; Golomb et al., 2008; Golomb et al., 2014; Golomb & Kanwisher, 2012;

Hannula et al., 2010; Kragel & Voss, 2022; Liu et al., 2017; Ryan et al., 2020; Voss et al., 2017; Wynn et al., 2019; Wynn et al., 2020; Yeung et al., 2017; Yeung et al., 2019). For example, eye tracking can be used to examine the errors made by participants as they switch their gaze between target and nontarget object features (i.e., object-to-location binding errors). Previous eye-tracking studies have found that amnesic patients as well as older adults at risk for Alzheimer's disease show altered gaze behavior compared with healthy controls, which was interpreted as interference between the contents of memory with the perceived object lure (Barense et al., 2012; Erez et al., 2013; Yeung et al., 2017; Yeung et al., 2019). Our online mouse-tracking study extends this work, showing that memory load (i.e., the number of objects held in working memory) influences object reconstruction (Figs. 5, 6) and increases object-to-location binding errors (Fig. 7). Future studies could more directly examine how age and risk for cognitive decline may interact with a continuous mouse-tracking measure during working memory. One possibility could be to examine both the binding errors that occur at the object-level (as was done in the present study; Fig. 7), but also at the feature-level (i.e., the individual nontarget shapes and colors bound to spatial locations). Such an approach could compare the similarities and differences between eye-tracking and mouse-tracking more directly, identifying important factors like the type of binding error (i.e., object-to-location or feature-to-location), time, and search variability.

Another line of future investigation may be to consider whether aspects of our mouse-tracking work are applicable to navigation tasks derived from the Morris water maze (Morris, 1984), as well as to the literature on spatial schemas more generally. The Morris water maze is a paradigm widely used to investigate rodents' navigational abilities, in which rodents are placed in a water tank and learn to navigate to the location of a hidden platform on which they can stand. This paradigm, combined with the analysis of search trajectories, has recently been popularized for studying the formation of spatial schemas in rodents (Richards et al., 2014), and variations of schema learning tasks have been studied in humans used computer-based versions (Antony et al., 2022; Cockcroft et al., 2022; Tompary et al., 2020). Future work could explore the similarities and differences between the search paths of mouse cursors towards target locations and the paths taken by actual mice towards physical platforms. Visual inspection of our human participant mouse path trajectories shows that they resemble actual mouse behavior (Fig. 7a): When a nontarget mouse hover occurs during a trial, participants generally make a detour (they "check out a location") before continuing to navigate towards the target object. These trajectory errors, like the false memories for platform locations on a Morris water maze, yield meaningful information about participants' internal representations of the target and nontarget objects.

A notable difference, however, is that previous water-maze studies are analyzed over physical space (with actual mice in a water maze) or 2D coordinate space (with abstract human tasks), whereas our study examines *shape–color feature space*. That is, our mouse-tracking approach could be used to extend existing research on the formation and influence of spatial schemas into an investigation of how visual features are integrated into analogous higher-order integrated representations. To give one example, an analysis of statistical divergence is often used in the spatial schema literature, in order to examine how closely rodents' or participants' search trajectories fit a learned distribution of spatial locations over time (Antony et al., 2022; Richards et al., 2014). A closer statistical fit is taken as evidence that navigation is being guided or influenced by aggregate information about spatial locations—that is, by spatial schemas. While we did not conduct this analysis here, a future study could investigate the analogous phenomenon over shape–color space, where participants' search trajectories map onto a statistical distribution of objects defined by shape–color dimensions. Similarly, key findings on the structure of spatial information, such as the role of grid and place cells in representing space (Bellmund et al., 2018; Moser et al., 2015) can also be investigated in the domain of shape–color object space. For instance, might some visual neurons encode conjunctions in a learned shape–color space, just as place cells do in coordinate space? Alternatively, it may be the case that "navigation" through physical space and shape–color object space relies on a common coordinate system (i.e., a cognitive map; de Cothi et al., 2022; Epstein et al., 2017; Theves et al., 2019; Whittington et al., 2022). For these reasons, applying trajectory analyses in the domain of continuous object space may yield new insights into the neural basis of feature integration across generalized task contexts during working memory.

Comparisons with existing online testing approaches

In developing our task, we created an open-source pipeline for developing online tasks in Python (see https://github.com/james-y-yuan/executable-pipeline for details, including additional timing measures). Although downloadable executables have been described at least over 2 decades ago (see Hewson et al., 1996), our work lays out the concrete implementation of this idea, providing open-source code in a freely accessible programming language. Moreover, we compared the data reliability of participants when tested individually and concurrently over virtual conference calls (i.e., over Zoom), replicating memory results within 4% of values (see Fig. 4). Importantly, we highlight that concurrent testing may be a straightforward method to increase testing efficiency while maintaining high data quality—even with existing online approaches (e.g., online experiments built though

JavaScript; Anwyl-Irvine et al., 2021; de Leeuw, 2015; de Leeuw & Motz, 2016). We suggest downloadable executables complement existing browser-based approaches to online testing; as with any tool, researchers must carefully consider the benefits and weaknesses of applying a given tool to their specific research question.

In our study, downloadable executables greatly simplified transitioning a mouse-tracking task already built for in-person experimentation into a format amenable to online testing. However, one possible limitation of using downloadable executables for online testing is their potential for security concerns. For example, researchers can hypothetically access aspects of the participant computer, like keyboard responses outside of the experiment or other malicious actions. We note that this potential security concern also exists with online testing using browser-based experiments, such that a researcher could in theory ask participants to click on a malicious link during an online task. In our study, we explicitly minimized these possible security concerns, as we recorded information from only the position of the mouse cursor and select keys during the experiment (see Fig. 2). The data was further anonymized and stored on our secure lab server, and we digitally signed the executable with detailed information from our laboratory. Finally, participants downloaded the executable from our lab website and completed informed consent directly with the experimenters in a virtual conference room through video call.

Importantly, we collected online data from 30 participants in just six testing sessions in the concurrent testing condition while maintaining high data quality (Figs. 3, 4). In a typical online experiment, participants often complete tasks with no oversight. By contrast, we speculate that the social pressure of having the experimenter present with other participants during a synchronous virtual conference call may have increased the likelihood that participants will focus on their task. Future studies could explore different experimental setups, such as a version of our online approach where participants are tested with their video camera on. Just as students in an online university course report feeling increased engagement, accountability, and connectedness when their cameras are on as opposed to off (Schwenck & Pryor, 2021), participants may similarly complete an experimental task in a more engaged manner with their cameras on. Though such an approach may increase conscientiousness in participants, future studies will also need to address privacy concerns, including the possibility that others can view their living circumstances. Nevertheless, our results suggest that testing many participants in the same virtual conference call (i.e., over Zoom, with or without video cameras on) may be a straightforward method to collect high quality data, even for approaches that do not involve downloadable executables (e.g., online browser-based experiments; Anwyl-Irvine et al., 2021; Sauter et al., 2020).

Conclusion

Altogether, our results provide evidence that capacity limits in visual working memory may arise partly from resolving—or failing to resolve—interference during object reconstruction. We found two sources of between-item interference during object reconstruction that increased with memory load: (i) the similarity match between the contents of working memory with what is currently being viewed (Figs. 5, 6) and (ii) object-to-location binding errors from previously studied but currently irrelevant lure objects (Fig. 7). Indeed, our trajectory analyses quantify object reconstruction behavior during the entire test phase of a simultaneous reconstruction task, converting what would typically be a dependent variable with a single observation each trial (i.e., the final response) to a dependent variable with hundreds of observations each trial at millisecond temporal resolution (i.e., time-series data). For these reasons, analyzing only the final response on a continuous memory task may fail to capture important aspects of behavior, akin to analyzing only the final gaze position during eye tracking or only the final position of a rodent during spatial learning. Potential future work could explore whether factors such as motor movements (Marteniuk et al., 1987), value-based choice (de Martino & Cortese, 2022), attention (Dowd & Golomb, 2019; Olivers et al., 2006), working memory distortions (Chunharas et al., 2022), ensemble coding (Whitney & Yamanashi Leib, 2018), or subjective similarity between the to-be-remembered objects (Lin & Luck, 2009) influence object reconstruction behavior beyond the effects we present here. In conclusion, our results show that measuring mouse-tracking behavior permits a more complete understanding of the internal representations leading up to the final response on a cognitive task, providing new insight into the dynamic interplay between working memory and perception.

Declarations

Conflicts of interest None declared.

Ethics approval REB approval was received from the University of Toronto (Protocol 38856).

Consent to participate Informed consent to participate was obtained for all individual participants included in the study.

Consent to publish Informed consent to publish was obtained for all individual participants included in the study.

References

Antony, J. W., Stiver, C. A., Graves, K. N., Osborne, J., Turk-Browne, N. B., & Bennion, K. A. (2022). Spatial gist extraction during human memory consolidation. *Journal of Experimental psychology: Learning, Memory, and Cognition, 48*(7), 929–941. https://doi.org/10.1037/xlm0000894

Anwyl-Irvine, A., Dalmaijer, E. S., Hodges, N., & Evershed, J. K. (2021). Realistic precision and accuracy of online experiment platforms, web browsers, and devices. *Behavior Research Methods, 53*(4), 1407–1425. https://doi.org/10.3758/s13428-020-01501-5

Baker, D. H., Vilidaite, G., Lygo, F. A., Smith, A. K., Flack, T. R., Gouws, A. D., & Andrews, T. J. (2020). Power contours: Optimising sample size and precision in experimental psychology and human neuroscience. *Psychological Methods*. https://doi.org/10.1037/met0000337

Barense, M. D., Groen, I. I., Lee, A. C., Yeung, L. K., Brady, S. M., Gregori, M., Kapur, N., Bussey, T. J., Saksida, L. M., & Henson, R. N. (2012). Intact memory for irrelevant information impairs perception in amnesia. *Neuron, 75*(1), 157–167. https://doi.org/10.1016/j.neuron.2012.05.014

Bartsch, L. M., & Oberauer, K. (2023). The contribution of episodic long-term memory to working memory for bindings. *Cognition, 231*, Article 105330. https://doi.org/10.1016/j.cognition.2022.105330

Bays, P. (2016). Evaluating and excluding swap errors in analogue tests of working memory. *Scientific Reports, 6*, Article 19203. https://doi.org/10.1038/srep19203

Bays, P. M., Catalao, R. F., & Husain, M. (2009). The precision of visual working memory is set by allocation of a shared resource. *Journal of Vision, 9*(10), 1–11. https://doi.org/10.1167/9.10.7

Bellmund, J., Gärdenfors, P., Moser, E. I., & Doeller, C. F. (2018). Navigating cognition: Spatial codes for human thinking. *Science, 362*(6415), eaat6766. https://doi.org/10.1126/science.aat6766

Bornstein, A. M., Aly, M., Feng, S. F., Turk-Browne, N. B., Norman, K. A., & Cohen, J. D. (2017). Perceptual decisions result from the continuous accumulation of memory and sensory evidence. *bioRxiv*. https://doi.org/10.1101/186817

Chunharas, C., Rademaker, R. L., Brady, T. F., & Serences, J. T. (2022). An adaptive perspective on visual working memory distortions. *Journal of Experimental Psychology: General*. Advance online publication. https://doi.org/10.1037/xge0001191

Cockcroft, J. P., Berens, S. C., Gaskell, M. G., & Horner, A. J. (2022). Schematic information influences memory and generalisation behaviour for schema-relevant and -irrelevant information. *Cognition, 227*, Article 105203.

Cooper, R. A., & Ritchey, M. (2019). Cortico-hippocampal network connections support the multidimensional quality of episodic memory. *eLife, 8*, Article e45591. https://doi.org/10.7554/eLife.45591

Cowan, N. (2001). The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behavioral and Brain Sciences, 24*(1), 87–185. https://doi.org/10.1017/s0140525x01003922

Cowan, N. (2010). The magical mystery four: How is working memory capacity limited, and why? *Current Directions in Psychological Science, 19*(1), 51–57. https://doi.org/10.1177/0963721409359277

Cowell, R. A., Barense, M. D., & Sadil, P. S. (2019). A roadmap for understanding memory: Decomposing cognitive processes into operations and representations. *eNeuro, 6*(4). https://doi.org/10.1523/ENEURO.0122-19.2019

de Cothi, W., Nyberg, N., Griesbauer, E.-M., Ghanamé, C., Zisch, F., Lefort, J. M., Fletcher, L., Newton, C., Renaudineau, S., Bendor, D., Grieves, R., Duvelle, É., Barry, C., & Spiers, H. J. (2022). Predictive maps in rats and humans for spatial navigation. *Current Biology*. https://doi.org/10.1016/j.cub.2022.06.090

de Leeuw, J. R. (2015). jsPsych: A JavaScript library for creating behavioral experiments in a Web browser. *Behavior Research Methods, 47*(1), 1–12. https://doi.org/10.3758/s13428-014-0458-y

de Leeuw, J. R., & Motz, B. A. (2016). Psychophysics in a Web browser? Comparing response times collected with JavaScript and Psychophysics Toolbox in a visual search task. *Behavior Research Methods, 48*, 1–12. https://doi.org/10.3758/s13428-015-0567-2

de Martino, B., & Cortese, A. (2022). Goals, usefulness, and abstraction in value-based choice. *Trends in Cognitive Science*. https://doi.org/10.1016/j.tics.2022.11.001

Dowd, E. W., & Golomb, J. D. (2019). Object-feature binding survives dynamic shifts of spatial attention. *Psychological Science, 30*(3), 343–361. https://doi.org/10.1177/0956797618818481

Endress, A. D., & Szabó, S. (2017). Interference and memory capacity limitations. *Psychological Review, 124*(5), 551–571. https://doi.org/10.1037/rev0000071

Epstein, R., Patai, E., Julian, J., & Spiers, H. J. (2017). The cognitive map in humans: Spatial navigation and beyond. *Nature Neuroscience, 20*, 1504–1513. https://doi.org/10.1038/nn.4656

Erez, J., Lee, A. C., & Barense, M. D. (2013). It does not look odd to me: Perceptual impairments and eye movements in amnesic patients with medial temporal lobe damage. *Neuropsychologia, 51*(1), 168–180. https://doi.org/10.1016/j.neuropsychologia.2012.11.003

Evans, N. J., & Wagenmakers, E.-J. (2020). Evidence accumulation models: Current limitations and future directions. *The Quantitative Methods for. Psychology, 16*(2), 73–90. https://doi.org/10.20982/tqmp.16.2.p073

Fradkin, I., & Eldar, E. (2022). Accumulating evidence for myriad alternatives: Modeling the generation of free association. *Psychological Review*. Advance online publication. https://doi.org/10.1037/rev0000397

Golomb, J. D., & Kanwisher, N. (2012). Retinotopic memory is more precise than spatiotopic memory. *Proceedings of the National Academy of Sciences, 109*(5), 1796–1801. https://doi.org/10.1073/pnas.1113168109

Golomb, J. D., Chun, M. M., & Mazer, J. A. (2008). The native coordinate system of spatial attention is retinotopic. *Journal of Neuroscience, 28*(42), 10654–10662. https://doi.org/10.1523/JNEUROSCI.2525-08.2008

Golomb, J. D., L'Heureux, Z. E., & Kanwisher, N. (2014). Feature-binding errors after eye movements and shifts of attention. *Psychological Science, 25*(5), 1067–1078. https://doi.org/10.1177/0956797614522068

Hannula, D. E., Althoff, R. R., Warren, D. E., Riggs, L., Cohen, N. J., & Ryan, J. D. (2010). Worth a glance: Using eye movements to investigate the cognitive neuroscience of memory. *Frontiers in Human Neuroscience, 4*, Article 166. https://doi.org/10.3389/fnhum.2010.00166

Hao, Y., Li, X., Zhang, H., & Ku, Y. (2021). Free-recall benefit, inhomogeneity and between-item interference in working memory. *Cognition, 214*, Article 104739. https://doi.org/10.1016/j.cognition.2021.104739

Hedayati, S., O'Donnell, R. E., & Wyble, B. (2022). A model of working memory for latent representations. *Nature Human Behaviour, 6*(5), 709–719. https://doi.org/10.1038/s41562-021-01264-9

Henderson, M. M., Rademaker, R. L., & Serences, J. T. (2022). Flexible utilization of spatial- and motor-based codes for the storage of visuo-spatial information. *eLife, 11*, Article e75688. https://doi.org/10.7554/eLife.75688

Hewson, C. M., Laurent, D., & Vogel, C. M. (1996). Proper methodologies for psychological and sociological studies conducted via the Internet. *Behavior Research Methods, 28*, 186–191. https://doi.org/10.3758/BF03204763

Hubel, D. H., & Wiesel, T. N. (1962). Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *The Journal of Physiology, 160*(1), 106–154. https://doi.org/10.1113/jphysiol.1962.sp006837

Kragel, J. E., & Voss, J. L. (2022). Looking for the neural basis of memory. *Trends in Cognitive Sciences, 26*(1), 53–65. https://doi.org/10.1016/j.tics.2021.10.010

Krueger, P. M., van Vugt, M. K., Simen, P., Nystrom, L., Holmes, P., & Cohen, J. D. (2017). Evidence accumulation detected in BOLD signal using slow perceptual decision making. *Journal of Neuroscience Methods, 281*, 21–32. https://doi.org/10.1016/j.jneumeth.2017.01.012

Li, M., Liu, F., Juusola, M., & Tang, S. (2014). Perceptual color map in macaque visual area V4. *Journal of Neuroscience, 34*(1), 202–217. https://doi.org/10.1523/JNEUROSCI.4549-12.2014

Li, A. Y., Liang, J. C., Lee, A., & 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

Li, A. Y., Fukuda, K., & Barense, M. D. (2022). Independent features form integrated objects: Using a novel shape–color "conjunction task" to reconstruct memory resolution for multiple object features simultaneously. *Cognition, 223*, 1–19. https://doi.org/10.1016/j.cognition.2022.105024

Liang, J. C., Erez, J., Zhang, F., Cusack, R., & Barense, M. D. (2020). Experience transforms conjunctive object representations: Neural evidence for unitization after visual expertise. *Cerebral Cortex, 30*(5), 2721–2739. https://doi.org/10.1093/cercor/bhz250

Lin, P. H., & Luck, S. J. (2009). The influence of similarity on visual working memory representations. *Visual Cognition, 17*(3), 356–372. https://doi.org/10.1080/13506280701766313

Liu, Z. X., Shen, K., Olsen, R. K., & Ryan, J. D. (2017). Visual sampling predicts hippocampal activity. *Journal of Neuroscience, 37*(3), 599–609. https://doi.org/10.1523/JNEUROSCI.2610-16.2016

Luck, S. J., & Vogel, E. K. (1997). The capacity of visual working memory for features and conjunctions. *Nature, 390*(6657), 279–281. https://doi.org/10.1038/36846

Ma, W. J., Husain, M., & Bays, P. M. (2014). Changing concepts of working memory. *Nature Neuroscience, 17*(3), 347–356. https://doi.org/10.1038/nn.3655

Magezi, D. A. (2015). Linear mixed-effects models for within-participant psychology experiments: An introductory tutorial and free, graphical user interface (LMMgui). *Frontiers in Psychology, 6*, 2. https://doi.org/10.3389/fpsyg.2015.00002

Marteniuk, R. G., MacKenzie, C. L., Jeannerod, M., Athenes, S., & Dugas, C. (1987). Constraints on human arm movement trajectories. *Canadian Journal of Psychology, 41*(3), 365–378. https://doi.org/10.1037/h0084157

Martin, C. B., Douglas, D., Newsome, R. N., Man, L. L., & Barense, M. D. (2018). Integrative and distinctive coding of visual and conceptual object features in the ventral visual stream. *eLife, 7*, Article e31873. https://doi.org/10.7554/eLife.31873

Morris, R. (1984). Developments of a water-maze procedure for studying spatial learning in the rat. *Journal of Neuroscience Methods, 11*(1), 47–60. https://doi.org/10.1016/0165-0270(84)90007-4

Moser, M. B., Rowland, D. C., & Moser, E. I. (2015). Place cells, grid cells, and memory. *Cold Spring Harbor Perspectives on Biology, 7*(2), Article a021808. https://doi.org/10.1101/cshperspect.a021808

Oberauer, K. (2019). Working memory capacity limits memory for bindings. *Journal of Cognition, 2*(1), Article 40. https://doi.org/10.5334/joc.86

Oberauer, K. (2021). Measurement models for visual working memory—A factorial model comparison. *Psychological Review*. https://doi.org/10.1037/rev0000328

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

Olivers, C. N., Meijer, F., & Theeuwes, J. (2006). Feature-based memory-driven attentional capture: Visual working memory content affects visual attention. *Journal of Experimental Psychology: Human Perception and Performance, 32*(5), 1243–1265. https://doi.org/10.1037/0096-1523.32.5.1243

Park, H.-B., & Zhang, W. (2022). Trial-by-trial mouse trajectory predicts variance in precision across working memory representations: A critical reanalysis of Hao et al. (2021). *Psychonomic Bulletin & Review*. https://doi.org/10.3758/s13423-022-02128-7

Parra, M. A., Abrahams, S., Fabi, K., Logie, R., Luzzi, S., & Della Sala, S. (2009). Short-term memory binding deficits in Alzheimer's disease. *Brain, 132*(4), 1057–1066. https://doi.org/10.1093/brain/awp036

Rademaker, R. L., Park, Y. E., Sack, A. T., & Tong, F. (2018). Evidence of gradual loss of precision for simple features and complex objects in visual working memory. *Journal of Experimental Psychology: Human Perception and Performance, 44*(6), 925–940. https://doi.org/10.1037/xhp0000491

Richards, B. A., Xia, F., Santoro, A., Husse, J., Woodin, M. A., Josselyn, S. A., & Frankland, P. W. (2014). Patterns across multiple memories are identified over time. *Nature Neuroscience, 17*(7), 981–986. https://doi.org/10.1038/nn.3736

Ryan, J. D., Althoff, R. R., Whitlow, S., & Cohen, N. J. (2000). Amnesia is a deficit in relational memory. *Psychological Science, 11*(6), 454–461. https://doi.org/10.1111/1467-9280.00288

Ryan, J. D., Leung, G., Turk-Browne, N. B., & Hasher, L. (2007). Assessment of age-related changes in inhibition and binding using eye movement monitoring. *Psychology and Aging, 22*(2), 239–250. https://doi.org/10.1037/0882-7974.22.2.239

Ryan, J. D., Shen, K., & Liu, Z. X. (2020). The intersection between the oculomotor and hippocampal memory systems: empirical developments and clinical implications. *Annals of the New York Academy of Sciences, 1464*(1), 115–141. https://doi.org/10.1111/nyas.14256

Sauter, M., Draschkow, D., & Mack, W. (2020). Building, hosting and recruiting: A brief introduction to running behavioral experiments online. *Brain Sciences, 10*(4), Article 251. https://doi.org/10.3390/brainsci10040251

Schneegans, S., & Bays, P. M. (2017). Neural architecture for feature binding in visual working memory. *The Journal of Neuroscience, 37*(14), 3913–3925. https://doi.org/10.1523/JNEUROSCI.3493-16.2017

Schneegans, S., McMaster, J. M. V., & Bays, P. M. (2022). Role of time in binding features in visual working memory. *Psychological Review*. Advance online publication. https://doi.org/10.1037/rev0000331

Schurgin, M. W., Wixted, J. T., & Brady, T. F. (2020). Psychophysical scaling reveals a unified theory of visual memory strength. *Nature Human Behaviour, 4*(11), 1156–1172. https://doi.org/10.1038/s41562-020-00938-0

Schwenck, C. M., & Pryor, J. D. (2021). Student perspectives on camera usage to engage and connect in foundational education classes: It's time to turn your cameras on. *International Journal of Educational Research Open, 2*, Article 100079. https://doi.org/10.1016/j.ijedro.2021.100079

Shenhav, A., Straccia, M. A., Musslick, S., Cohen, J. D., & Botvinick, M. M. (2018). Dissociable neural mechanisms track evidence accumulation for selection of attention versus action. *Nature. Communications, 9*, Article 2485. https://doi.org/10.1038/s41467-018-04841-1

Shepherdson, P., Hell, L., & Oberauer, K. (2022). How does visual working memory solve the binding problem? *Journal of Experimental Psychology: Human Perception and Performance, 48*(10), 1137–1152. https://doi.org/10.1037/xhp0001044

Shipstead, Z., & Engle, R. W. (2013). Interference within the focus of attention: Working memory tasks reflect more than temporary maintenance. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 39*(1), 277–289. https://doi.org/10.1037/a0028467

Sone, H., Kang, M. S., Li, A. Y., Tsubomi, H., & Fukuda, K. (2021). Simultaneous estimation procedure reveals the object-based, but not space-based, dependence of visual working memory representations. *Cognition, 209*, Article 104579. https://doi.org/10.1016/j.cognition.2020.104579

Sprague, T., & Serences, J. (2013). Attention modulates spatial priority maps in the human occipital, parietal and frontal cortices. *Nature Neuroscience, 16*, 1879–1887. https://doi.org/10.1038/nn.3574

Sutterer, D. W., & Awh, E. (2016). Retrieval practice enhances the accessibility but not the quality of memory. *Psychonomic Bulletin & Review, 23*(3), 831–841. https://doi.org/10.3758/s13423-015-0937-x

Swan, G., & Wyble, B. (2014). The binding pool: A model of shared neural resources for distinct items in visual working memory. *Attention, Perception, & Psychophysics, 76*(7), 2136–2157. https://doi.org/10.3758/s13414-014-0633-3

The Jamovi Project. (2021). *jamovi* (Version 1.6) [Computer Software]. https://www.jamovi.org

Theves, S., Fernandez, G., & Doeller, C. F. (2019). The hippocampus encodes distances in multidimensional feature space. *Current Biology, 29*(7), 1226–1231.e3. https://doi.org/10.1016/j.cub.2019.02.035

Thyer, W., Adam, K. C. S., Diaz, G. K., Velázquez Sánchez, I. N., Vogel, E. K., & Awh, E. (2022). Storage in visual working memory recruits a content-independent pointer system. *Psychological Science, 33*(10), 1680–1694. https://doi.org/10.1177/09567976221090923

Tompary, A., Zhou, W., & Davachi, L. (2020). Schematic memories develop quickly, but are not expressed unless necessary. *Scientific Reports, 10*, Article 16968. https://doi.org/10.1038/s41598-020-73952-x

Treisman, A. (1996). The binding problem. *Current Opinion in Neurobiology, 6*(2), 171–178. https://doi.org/10.1016/S0959-4388(96)80070-5

Treisman, A., & Gelade, G. (1980). A feature-integration theory of attention. *Cognitive Psychology, 12*, 97–136. https://doi.org/10.1016/0010-0285(80)90005-5

Voss, J. L., Bridge, D. J., Cohen, N. J., & Walker, J. A. (2017). A closer look at the hippocampus and memory. *Trends in Cognitive Sciences, 21*(8), 577–588. https://doi.org/10.1016/j.tics.2017.05.008

Whitney, D., & Yamanashi Leib, A. (2018). Ensemble perception. *Annual Review of Psychology, 69*, 105–129. https://doi.org/10.1146/annurev-psych-010416-044232

Whittington, J. C. R., McCaffary, D., Bakermans, J. J. W., & Behrens, T. E. J. (2022). How to build a cognitive map. *Nature Neuroscience, 25*, 1257–1272. https://doi.org/10.1038/s41593-022-01153-y

Woodworth, R. S. (1899). Accuracy of voluntary movement. *The Psychological Review: Monograph Supplements, 3*(3), i–114. https://doi.org/10.1037/h0092992

Wu, Z., & Buckley, M. J. (2022). Prefrontal and medial temporal lobe cortical contributions to visual short-term memory. *Journal of Cognitive Neuroscience*, 1–17. Advance online publication.. https://doi.org/10.1162/jocn_a_01937

Wynn, J. S., Shen, K., & Ryan, J. D. (2019). Eye movements actively reinstate spatiotemporal mnemonic content. *Vision, 3*(2), 21. https://doi.org/10.3390/vision3020021

Wynn, J. S., Ryan, J. D., & Buchsbaum, B. R. (2020). Eye movements support behavioral pattern completion. *Proceedings of the National Academy of Sciences, 117*(11), 6246–6254. https://doi.org/10.1073/pnas.1917586117

Xu, Z., Adam, K. C. S., Fang, X., & Vogel, E. K. (2017). The reliability and stability of visual working memory capacity. *Behavior Research Methods*. https://doi.org/10.3758/s13428-017-0886-6

Yeung, L.-K., Ryan, J. D., Cowell, R. A., & Barense, M. D. (2013). Recognition memory impairments caused by false recognition of novel objects. *Journal of Experimental Psychology: General, 142*(4), 1384–1397. https://doi.org/10.1037/a0034021

Yeung, L. K., Olsen, R. K., Bild-Enkin, H. E. P., D'Angelo, M. C. D., Kacollja, A., McQuiggan, D. A., Keshabyan, A., Ryan, J. D., & Barense, M. D. (2017). Anterolateral entorhinal cortex volume predicted by altered intra-item configural processing. *Journal of Neuroscience, 37*(22), 5527–5538. https://doi.org/10.1523/JNEUROSCI.3664-16.2017

Yeung, L. K., Olsen, R. K., Hong, B., Mihajlovic, V., D'Angelo, M. C., Kacollja, A., Ryan, J. D., & Barense, M. D. (2019). Object-in-place memory predicted by anterolateral entorhinal cortex and parahippocampal cortex volume in older adults. *Journal of Cognitive Neuroscience, 31*(5), 711–729. https://doi.org/10.1162/jocn_a_01385

Yonelinas, A. P., Ranganath, C., Ekstrom, A. D., & Wiltgen, B. J. (2019). A contextual binding theory of episodic memory: systems consolidation reconsidered. *Nature Reviews Neuroscience, 20*, 364–375. https://doi.org/10.1038/s41583-019-0150-4

Zhang, W., & Luck, S. J. (2008). Discrete fixed-resolution representations in visual working memory. *Nature, 453*(7192), 233–235. https://doi.org/10.1038/nature06860

Zhang, W., & Luck, S. J. (2009). Sudden death and gradual decay in visual working memory. *Psychological Science, 20*(4), 423–428. https://doi.org/10.1111/j.1467-9280.2009.02322.x