Title

Viewing memory through the eyes: Recognition processes, semantic memory, and attention

Author

Ramey, Michelle

Publication Date

2021

Viewing memory through the eyes:
Recognition processes, semantic memory, and attention

By

MICHELLE M. RAMEY
DISSERTATION

Submitted in partial satisfaction of the requirements for the degree of

DOCTOR OF PHILOSOPHY

in

Psychology

in the

OFFICE OF GRADUATE STUDIES

of the

UNIVERSITY OF CALIFORNIA

DAVIS

Approved:

Andrew P. Yonelinas, Chair

John M. Henderson

Steve J. Luck

Committee in Charge

2021

Table of Contents

Abstract

What we attend to determines what we remember, and what we remember influences what we attend to. Despite the evidence for a strong bidirectional relationship between memory and visual attention, exactly *how* they are related is not well understood, particularly with respect to the different processes that underlie memory. This dissertation examines both 1) how episodic memory and its underlying processes guide attention, and interact with semantic memory to do so, and 2) how attention influences encoding for different episodic memory processes. Chapter 1 presents a new method to tease apart the processes underlying recognition memory—namely recollection, familiarity, and unconscious memory—to show that memory-driven improvements to search (i.e., contextual cueing) are driven by both recollection and unconscious memory. Importantly, these two memory processes contributed to search through two distinct patterns of eye movements: unconscious memory improved overall efficiency, whereas recollection was related to highly accurate first eye movements that were aimed in the direction of the target. Chapter 2 presents two newly developed eye movement measures and explores the relationship between episodic memory and eye movements during encoding and retrieval of scenes, revealing that familiarity strength has a robust relationship with attention during both encoding and retrieval. During encoding, distributing attention broadly across the scene predicted subsequently stronger familiarity, whereas during retrieval, focusing attention on fewer regions was related to stronger familiarity. Furthermore, revisiting regions between encoding and retrieval (i.e., *resampling*) predicted stronger familiarity as well. To determine how these episodic memory effects on attention might be modulated by semantic memory, Chapter 3 examines how attention to semantic information in a scene affects the relationship between familiarity strength and resampling. Despite the literature historically characterizing resampling as a purely episodic effect, semantic memory was in fact more predictive of resampling, and it exhibited potentially competitive interactions with familiarity strength such that the relationship between familiarity and resampling was weaker when participants attended more to semantic information. Lastly, Chapter 4 further explores the potential for competitive interactions between semantic memory and episodic memory by directly probing the spatial memory representation that is harnessed to guide search. Every episodic process examined (i.e., recollection, familiarity, and unconscious memory) contributed to participants' memory for target location and competed with semantic memory to do so, but only recollection eliminated spatial memory bias by semantic memory altogether. Together, these findings indicate that our attention is guided by a complex interplay of multiple types of memory, with the involvement of different episodic processes depending on both the task and the availability of semantic memory influences. More generally, these results suggest that many of the debates surrounding the interactions between memory and attention could potentially be resolved by examining the processes underlying memory, rather than treating it as a single construct.

Acknowledgements

First, I would like to thank Dr. Andrew Yonelinas and Dr. John Henderson for all of their guidance, direction, encouragement, and support. Thank you for teaching me how to turn my convoluted ideas into something that actually makes sense to other people. I would also like to thank Dr. Steve Luck for mentoring me in experimental design, EEG methods, and helping me think more deeply about theory.

Thank you also to my family for your support, encouragement, and always being understanding of graduate school getting in the way of normal life functions. And thank you to all my friends, old and new, who have helped me through the tough times and celebrated with me in the good ones.

Most of all, I would like to thank my husband, who made me feel so welcome and at home from the moment I set foot in Davis and has been a wonderful mentor, friend, and inspiration all five years. I wouldn’t be the scientist—or person—I am today if it weren’t for you. I will always fondly remember graduate school as the time you went from my incredible colleague to my even more incredible husband.

Chapter 1

Conscious and unconscious memory differentially impact attention: Eye movements, visual search, and recognition processes

Michelle M. Ramey[a,b,c*], Andrew P. Yonelinas[a,b], John M. Henderson[a,c]

[a] Department of Psychology, University of California, Davis, CA, USA.

[b] Center for Neuroscience, University of California, Davis, CA, USA.

[c] Center for Mind and Brain, University of California, Davis, CA, USA.

Published in *Cognition* (2019)

Author Note:

Funding: This work was supported by the National Eye Institute of the National Institutes of Health under Award Numbers R01EY025999 and R01EY027792. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Author contributions: Conceptualization, Methodology, Writing – Review & Editing, J.M.H., A.P.Y., M.M.R.; Funding acquisition, A.P.Y., J.M.H.; Formal Analysis, Investigation, Writing – Original Draft, Visualization, M.M.R.

Conflict of interest: Authors declare no conflict of interest.

Abstract

A hotly debated question is whether memory influences attention through conscious or unconscious processes. To address this controversy, we measured eye movements while participants searched repeated real-world scenes for embedded targets, and we assessed memory for each scene using confidence-based methods to isolate different states of subjective memory awareness. We found that memory-informed eye movements during visual search were predicted both by conscious recollection, which led to a highly precise first eye movement toward the remembered location, and by unconscious memory, which increased search efficiency by gradually directing the eyes toward the target throughout the search trial. In contrast, these eye movement measures were not influenced by familiarity-based memory (i.e., changes in subjective reports of memory strength). The results indicate that conscious recollection and unconscious memory can each play distinct and complementary roles in guiding attention to facilitate efficient extraction of visual information.

Keywords: Recognition; Contextual cueing; Visual search; Eyetracking; Memory; Implicit memory

Conscious and unconscious memory differentially impact attention: Eye movements, visual search, and recognition processes

How people move their eyes can provide important clues about the contents of their mind. For example, how the eyes move while viewing an object or scene can differ for previously encountered items compared to those that are novel, suggesting that eye movements can be used to reveal memory for prior experiences (Althoff & Cohen, 1999; Hannula, 2010; Henderson & Hollingworth, 2003; Hollingworth, Williams, & Henderson, 2001; Kafkas & Montaldi, 2012; Ryan, Hannula, & Cohen, 2007). Interestingly, there is evidence that eye movements may reveal influences of memory even when participants have no conscious awareness of that memory (Hannula, 2010; Hannula, Baym, Warren, & Cohen, 2012; Ryan, Althoff, Whitlow, & Cohen, 2000; Smith & Squire, 2017). Whether these effects are truly due to conscious or unconscious memory, however, is hotly contested, because other studies have found that eye movement differences between new and previously viewed items are limited to trials in which participants are aware of their memory for the items (Smith, Hopkins, & Squire, 2006; Smith & Squire, 2008).

The debate about conscious and unconscious influences on visual attention has been particularly lively in the literature on contextual cueing—a paradigm in which participants search for a target letter in an array of distractors (Chun & Jiang, 1998). Repetition of the array and target combination leads to performance improvements due to memory for the context (i.e., the array of distractors), as evidenced by a decrease in response time, a more accurate first eye movement toward the target location, and a decrease in the number of eye movements needed to reach the target (Peterson & Kramer, 2001). Improved search performance in this task is often attributed to unconscious memory, because when participants are subsequently asked if they recognize the studied arrays, they report little or no memory for those arrays (Chun, 2000; Chun & Jiang, 1998; Chun & Phelps, 1999; Johnson, Woodman, Braun, & Luck, 2007). However, because these recognition tests did not probe memory confidence and are typically based on small numbers of trials, the failure to find a relationship with conscious memory could be attributed to imprecise memory probes and insufficient power (Goujon, Didierjean, & Marmèche, 2007; Schlagbauer, Muller, Zehetleitner, & Geyer, 2012; Smyth & Shanks, 2008). Moreover, in contextual cueing studies using realistic scenes as contexts, participants report recognizing the scenes well above chance after the conclusion of the study (Brockmole & Henderson, 2006a, 2006b). It is therefore currently unknown whether the improvements in search performance from contextual cueing are due to conscious memory, or to unconscious memory. Furthermore, it is not known whether conscious and unconscious memory might be able to influence eye movements in different ways.

Current Research

In the present experiment, we investigated these possibilities by combining a contextual cueing paradigm with a confidence-based recognition method to isolate different states of subjective memory awareness, and used eyetracking to examine different patterns of eye movements during search. Participants first completed a learning phase where they searched realistic scenes for embedded target letters, indicating when they identified the letter (i.e., either a “T” or an “L”). During the subsequent test phase (see Figure 1), participants were presented with a mixture of old scenes (i.e., scenes that had been encountered in the earlier learning phase) and new scenes; for each scene, they first made a recognition judgment, and then searched for the target letter as they did in the learning phase. Memory awareness was measured by asking participants to rate memory confidence for each scene on a 6-point scale during the recognition judgment. For each of these test scenes, participants were told that if they could consciously recollect some qualitative aspect of the initial learning event, such as what they thought about when the scene was encountered earlier, they should respond “Recollect old (6);” otherwise, they rated their memory confidence by responding “I’m sure it’s old (5),” “Maybe it’s old (4),” “I don’t know (3),” “Maybe it’s new (2),” or “I’m sure it’s new (1).”

To isolate eye movement patterns related to unambiguously unconscious memory, we examined the scenes that participants were confident had not been studied (i.e., receiving a response of “I’m sure it’s new”), and in this way we excluded any scenes for which there was even a weak sense of conscious memory (i.e., the “maybe it’s old’, “I don’t know” and “maybe it’s new” trials). Conversely, to assess eye movement patterns that were related to unambiguously conscious memory, we examined the old scenes that were confidently recognized as studied and for which participants reported being able to retrieve specific details about the study event (i.e., “recollect old”). We also examined intermediate levels of memory confidence to determine if memory strength might be related to eye movement patterns in a manner similar to conscious or unconscious memory. There is evidence from behavioral and neural studies that intermediate-confidence recognition responses may rely on a familiarity process that is distinct from conscious recollection (e.g., Eichenbaum, Yonelinas, & Ranganath, 2007; Yonelinas, 2002, but also see Donaldson, 1996), but whether they are related to distinct types of eye movements is not yet clear (e.g., Kafkas & Montaldi, 2011, 2012; Sharot et al., 2008).

The influences of the above memory processes on contextual cueing performance were assessed in the present study using two eye movement measures, which have been shown to be sensitive to contextual cueing effects and are reflective of eye movement behavior at distinct points in the search process. To assess search-relevant behavior early in the trial, we measured the degree error of the first eye movement in a trial (i.e., *first saccade accuracy*) in terms of whether it was aimed toward the target, similar to the measure used by Peterson & Kramer (2001). To assess search processes that take place throughout the course of the trial, we measured *scanpath efficiency*—the efficiency of the overall search path—by dividing the total distance traveled by the shortest possible path to the target (Castelhano & Henderson, 2007). We expected that both of these eye movement measures of performance would be improved for scenes that were repeated compared to scenes that were new, and would thus reflect influences of memory.

Given the ambiguity of the existing literature regarding the role of conscious and unconscious memory in eye movement guidance, we did not have strong *a priori* predictions about how the eye movement measures would be related to reports of memory awareness. However, a number of different outcomes would be of theoretical interest. For example, the findings may show that eye movements are related selectively to either conscious or unconscious memory, which would be useful in resolving the debate about whether eye movements reflect conscious or unconscious memory processes. Conversely, both conscious and unconscious memory may influence each of the eye movement measures, suggesting that the influence of memory on eye movements may not clearly dissociate along the lines of conscious awareness. Finally, conscious and unconscious memory may be related to distinct types of eye movements, suggesting that conscious and unconscious memory processes may contribute to visual search in distinct ways.

Figure 1. Test phase trial. In each test phase trial, participants saw a preview of a scene without its target, then reported whether they remembered viewing the scene in the learning phase. They subsequently searched the scene again as they did in the learning phase. (A) The preview of the scene, which did not contain the target letter. (B) A schematic of the recognition judgment probe; participants were asked to rate their memory for the scenes on a 1-6 scale capturing varying levels of confidence. (Note that colors and spacing for "Recollect old" and "I'm sure it's new" responses are used for the sake of illustration; they were not visually unique from the other responses in the actual memory probe.) (C) The scene presented again for the subsequent search task; a zoomed in view of the target letter ("L") is provided for visualization purposes.

Method

Participants

Twenty-six undergraduate students were recruited from the participant pool at the University of California, Davis. Three participants did not meet criteria for high-quality eyetracking (i.e., at least 75% signal, such that the eyetracker was able to record the position of the eyes at least 75% of the time) and were excluded from analysis. Signal in the 23 participants included in the analyses was 93% on average. The included sample exceeded the sample size needed to detect the weakest effects previously reported in similar studies with 95% power. All participants had normal or corrected-to-normal vision and participated in exchange for course credit. Informed consent was obtained from all participants.

Materials

Stimuli. Stimuli consisted of 160 digital photographs of real-world indoor and outdoor scenes (Figure 1). Each scene contained a small grey "L" or "T" as the search target. The search targets were randomly placed using a custom MATLAB script to ensure an even distribution of target locations across scenes, with the center region (160x160px) and periphery of the scenes (outer 15%) excluded as possible target locations. Scenes were subsequently manually inspected, and if targets were not visible, they were moved to the nearest possible point at which they were visible. This procedure resulted in targets with x-coordinates ranging from 154px to 871px, and y-coordinates ranging from 115px to 653px. The initial scene set consisted of 300 scenes, and these scenes were subsequently normed for search difficulty; the final set of 160 scenes was produced by eliminating the easiest and most difficult scenes. For a given scene, the identity of the target letter was consistent and appeared in the same location across presentations.

Of the 160 scenes, 64 scenes were presented once in the learning phase (i.e., 1x scenes), 64 scenes were presented three times in the learning phase (i.e., 3x scenes), and 32 scenes served as lures in the test phase (i.e., new scenes). To minimize any stimulus effects, two counterbalances were used: assignment of scene to condition was counterbalanced such that each scene appeared in two of these three possible conditions (1x, 3x, new) across participants. That is, across participants, each scene appeared as a 1x and new scene, a 1x and 3x scene, or a 3x and new scene.

Apparatus. Participants sat 85cm away from the screen, such that scenes spanned approximately 25°x19° of visual angle at a resolution of 1024x768 pixels. Targets spanned approximately 0.27°x0.27° of visual angle. Eye movements were recorded using an SR research Eyelink 1000+ tower mount eyetracker, which sampled at 1000Hz. Eye movements were measured from the right eye, although vision was binocular, and a chin and forehead rest were used to prevent head movements. Participants' eye movements were tracked throughout the experiment.

Procedure

The first half of the experiment consisted of a learning phase, during which participants searched a series of scenes for their targets. Each scene was preceded by a fixation cross to ensure that participants were looking at the center of the screen upon scene presentation. Participants were asked to find the search target in each scene and use the keyboard to indicate whether it was an “L” or a “T”. Each trial was terminated at response, or after 20s without response. Participants were not aware that there would be a subsequent memory test and were not told that scenes would be repeated.

The learning phase was comprised of 128 unique scenes, with 64 scenes presented once (i.e. 1x scenes) and 64 scenes presented three times (i.e. 3x scenes), for a total of 256 trials. The 3x scenes were included to assess contextual cueing effects during the learning phase over repeated searches. Based on prior research, we anticipated near-ceiling memory performance for the 3x scenes, and therefore included the 1x scenes to generate a more even distribution of recognition responses for the purposes of test phase analyses by memory type. The scenes were randomly ordered throughout the learning phase with the constraint that at least 3 different scenes were interleaved between repeated presentations of a given scene. The participants were given short breaks every 50 trials and between the learning and test phases, and the eyetracker was re-calibrated after each break to prevent drift. The delay between the end of the learning phase and the start of the test phase instructions was approximately three minutes.

In the test phase (Figure 1), each trial started with a fixation cross, followed by a 400 ms preview of a scene without its target; the scene was either one that had been presented in the learning phase, or a new scene. Each preview was then replaced by a memory response screen. Participants were given as long as they needed to provide a memory response indicating whether or not they recognized the scene from the learning phase. Response options fell on a 1-5 and recollect scale (Yonelinas, 2002) made up of “sure new”, “maybe new”, “don’t know”, “maybe old”, “sure old”, and “recollect old”. Participants were instructed and tested on how to use this scale prior to beginning the test phase. A response of “recollect old” indicated that a participant could recall details of their experience of having seen the image in the learning phase. Examples given to participants included remembering an emotion they felt during prior exposure to the scene, and remembering ambient noise or sensations experienced while previously viewing the scene. The other responses fell on a continuous gradient ranging from no memory to strong memory for a scene. Importantly, participants were instructed that a response of “sure old” indicated high memory confidence comparable to that of “recollect old” responses, but without the additional episodic details.

After the memory probe in each trial, participants were shown another fixation cross followed by the same scene—this time with the search target included—and searched as they did in the learning phase. The test phase consisted of 160 trials (128 old scenes and 32 lures), with one trial for each unique scene.

Analysis

Statistical analyses were conducted using linear mixed effects models with crossed random effects of participant and image, which allowed us to harness trial-by-trial (i.e., within-subjects) data while controlling for individual differences and stimulus effects. The models were estimated using the lmerTest package in R (Kuznetsova, Brockhoff, & Christensen, 2017), and were fit using maximum likelihood. The degrees of freedom, estimated using the Satterthwaite approximation, and *t* values used were output by the linear mixed effects model for the variables of interest. Effect sizes were calculated as classical Cohen's d, as $\frac{2t}{\sqrt{df}}$ (Rosenthal & Rosnow, 1991). Trials with less than 75% signal were excluded from analysis (5% of trials).

The models for each analysis of the effects of memory on eye movements were specified by regressing the eye movement measure in question (i.e., first saccade accuracy, scanpath ratio, or delta distance) on a memory variable, which depended on the type of memory being assessed: conscious recollection, unconscious memory, or familiarity strength (see Table 1). In order to plot the data in a way that most directly reflects the analyses, Figures 2, 3a, and 4a are derived from the linear mixed effects models. As such, the plotted means are estimated marginal means, and the error bars represent the standard error of the estimated marginal means, both derived from the models. This allows for both within- and between-subjects variance to be taken into account, and controls for participant and stimulus effects. (However, the plots are similar when the raw data is plotted. See Appendix for more details.)

Table 1
Linear Mixed Effects Model Specifications for Each Test Phase Analysis

Fixed effect: Memory contrast	**Random effects**	**Included data**
Recollection: "Recollect" versus "Sure old"	Participant, image	All old scenes given responses of "recollect" or "sure old"
Unconscious memory: "Sure new" old scene versus "Sure new" new scene	Participant, image	All scenes, both old and new, given a "sure new" response
Familiarity strength: "Sure new" through "Sure old"	Participant, image	All old scenes except those given a "recollect" response

Note: In each model, the outcome was the eye movement measure of interest (i.e., first saccade accuracy, scanpath ratio, or delta distance). The fixed effect was the memory response given to a scene, or, in the case of unconscious memory, the old versus new status of the scene. Random effects were selected *a priori* to control for potential confounding influences of participant and scene, given the repeated measures design.

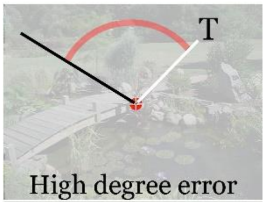
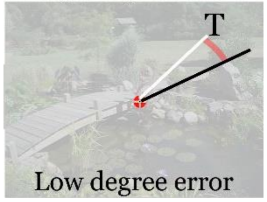
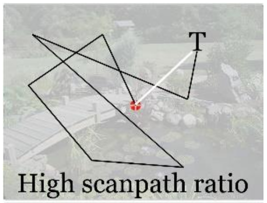
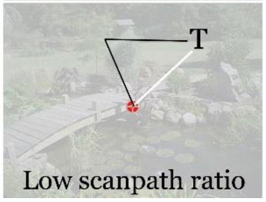
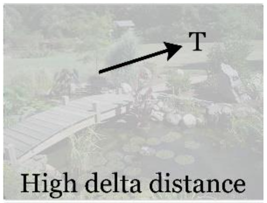
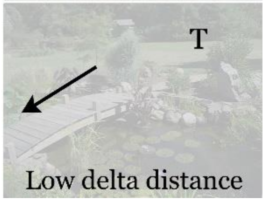
Results

Contextual Cueing Effects

Search speed. The learning phase included a mixture of scenes that were presented once, which we used to examine memory in the subsequent test phase, and scenes that were presented three times, which allowed us to examine contextual cueing effects prior to probing recognition memory. For the scenes that were presented three times in the learning phase, participants showed a significant increase in search speed across repeated presentations, $t(6410) = -16.84$, $p < .0001$, $d = -0.42$ (first presentation M=6030ms, third presentation M=3296ms), and repeated scenes had faster search speed than new scenes in the test phase, $t(2347) = -9.44$, $p < .0001$, $d = -0.39$, demonstrating the standard contextual cueing effect (Brockmole & Henderson, 2006a, 2006b).

Table 2

Descriptions of Eye Movement Measures and Terms Used

Eye movement term	Description	
Saccade	A movement of the eyes between two locations. Represented as black lines in the following figures.	
First saccade accuracy	The accuracy with which the first saccade in a trial was directed towards the target. Calculated as the angular degree error between the vector defined by the first saccade (from the center starting point), and the vector defined by the direct path to the target from the center starting point. A higher degree error value indicates that the first saccade was less accurate. One degree error value was obtained per trial.	 
Scanpath efficiency	The efficiency with which the eyes traveled to the target throughout the course of the trial, defined as the ratio of the eyes' actual path to the shortest possible path to the target. Calculated as the scanpath ratio, which is the total distance traveled by the eyes in a trial divided by the path directly from the center starting point to the target. A value of one indicates ideal performance, such that the actual path was the same length as the ideal path, and higher scanpath ratio values indicate poorer performance. One scanpath ratio value was obtained per trial.	 
Delta distance	The extent to which each saccade brought the eyes closer to the target. Calculated as the change in distance from the target that resulted from a saccade, given in pixels. A positive value indicates that the landing location of a saccade was closer to the target than the starting location of the saccade. One delta distance value was obtained per saccade per trial. The arrows represent saccades, with direction of the saccade indicated.	 

Note: Figures are included to provide examples of high and low values for each measure. In these figures, each black line (and arrow) represents a saccade. The white line represents ideal performance, which serves as the comparison point in calculating first saccade accuracy and scanpath efficiency.

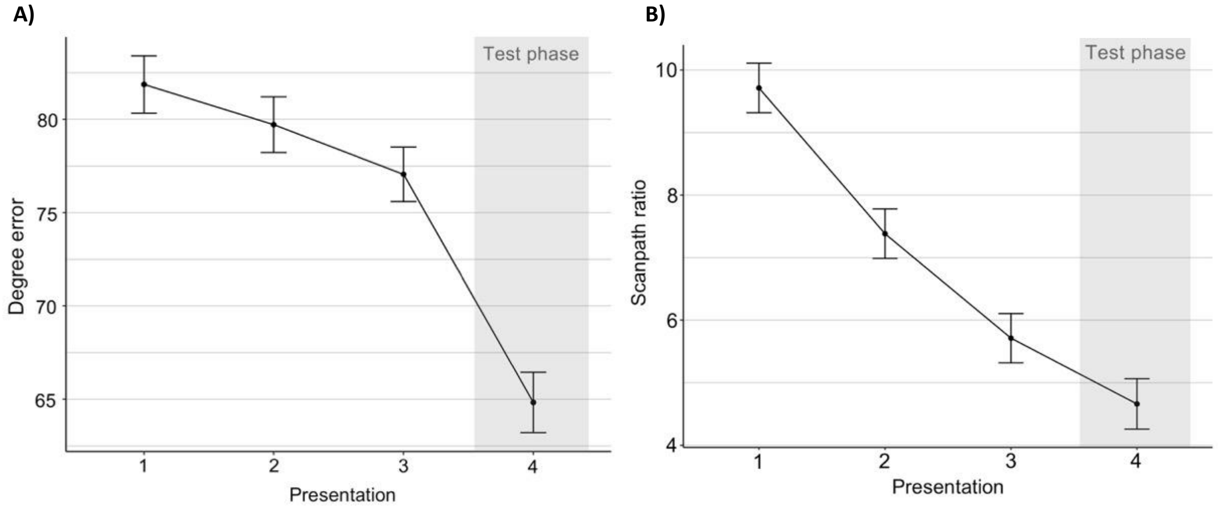


Figure 2. Improvement in eye movement measures for scenes presented three times in the learning phase, over the course of the learning phase (presentations 1-3) and the test phase (presentation 4). Least-squares means controlling for participant and image are plotted, and the error bars represent the standard error of these estimated means from the model. (A) First saccade accuracy. (B) Scanpath efficiency.

Eye movements. To investigate the behavior of the first eye movement made in the search process, first saccade accuracy was calculated as the degree error between the vector defined by the first saccade and the vector defined by the ideal path to the target from the central fixation cross (Table 2). Lower values indicate lower degree error and therefore better performance, such that the first eye movement was aimed more directly towards the target. First saccade accuracy improved across presentations in the learning phase (across presentations 1-3), *t*(3834)= -2.37, *p*=.018, *d*= -0.08, and throughout the experiment as a whole (across presentations 1-4), *t*(4971)= -7.88, *p*<.0001, *d*= -0.22 (Figure 2a).

To investigate eye movement behavior throughout the search period, we quantified scanpath efficiency over the course of the trial as the *scanpath ratio* (Castelhano & Henderson, 2007), the ratio of the observed scanpath to the most direct possible path to the target (Table 2). Specifically, we defined scanpath ratio as the total distance travelled by the eyes in the course of finding the target (calculated by summing the length of saccades) divided by the length of the direct path from the starting point to the target. As such, a scanpath ratio of 1 indicates ideal performance, such that the path taken by the eyes was the same length as the ideal path, whereas higher scanpath ratio values indicate less efficient paths. Scanpath ratio also improved over the course of the learning phase, $t(3379)= -14.73$, $p<.0001$, $d= -0.51$, and throughout the entire experiment, $t(4317)= -18.65$, $p<.0001$, $d= -0.57$ (Figure 2b).

Both first saccade accuracy and scanpath efficiency were also significantly better for repeated scenes than new scenes in the test phase (first saccade accuracy: $t(1515)= -5.87$, $p<.0001$, $d= -0.30$; scanpath efficiency: $t(2131)= -9.37$, $p<.0001$, $d= -0.41$). Together, these analyses suggest that the two eye movement measures improved as a result of contextual cueing, and that these effects were observed even prior to the test phase in which recognition judgments were required.

Behavior During the Test Phase

Memory accuracy. The percentage of scenes receiving a recognition confidence response corresponding to "recollect," "sure old," "maybe old," "don't know," "maybe new," and "sure new," respectively, were 41%, 34%, 10%, 6%, 5%, 4%, for scenes that were presented three times during the learning phase; 11%, 24%, 14%, 12%, 19%, 20% for scenes presented once during the learning phase; and 0.5%, 2.5%, 5%, 13%, 25%, 54% for new scenes. These results indicate that participants recognized more of the old than new scenes, and they used the full range of response confidence ratings. Only the scenes presented once in the learning phase and the new scenes were used in each of the following test phase analyses.

First saccade accuracy. Figure 3a presents test phase first saccade accuracy, in terms of degree error, as a function of memory response and type of scene (i.e., old versus new). To determine whether first saccade accuracy was influenced by recollection, we examined the test phase eye movement data from old scenes that were endorsed as “recollect old” (i.e., recollection for the scene) and compared this to the eye movement data from old scenes endorsed as “sure old” (i.e., high-confidence familiarity with the scene) (Table 1). We found that conscious recollection of a scene drove the first saccade more directly toward the target than did high-confidence familiarity with a scene, *t*(327)= -3.6, *p*<.001, *d*= -0.40. This result indicates that conscious recollection improved the accuracy of the first eye movement, such that the first saccade was more likely to be aimed toward the target in scenes that were recollected.

To determine whether first saccade accuracy was influenced by unconscious memory, we examined the test phase data from old scenes that were endorsed as “sure new” and compared them to the data from new scenes that were endorsed as “sure new” (Table 1). This strict criterion for unconscious memory (i.e., only considering scenes that were endorsed as “sure new” rather than all misses) ensured that none of the scenes used in the unconscious memory contrast were contaminated by conscious recollection or familiarity, and that the scenes differed only in terms of whether or not the participant had seen them previously. Put differently, we compared first saccade accuracy between high-confidence misses and high-confidence correct rejections. First saccade accuracy did not differ between the new scenes and the old scenes that participants were confident they had not seen before, *t*(264)= -1.26, *p*=.21, *d*= -0.16, indicating that unconscious memory did not significantly improve first saccade accuracy. Additionally, Bayesian analysis provided substantial evidence for the null hypothesis, suggesting that the numerically better first saccade accuracy for previously viewed scenes was almost five times more likely to be explained by chance than by unconscious memory (BF_{10}= .22).[1]

---

[1] By convention, a BF_{10} < .33 indicates substantial evidence for the null hypothesis (Jeffreys, 1961).

Lastly, to determine whether first saccade accuracy was influenced by familiarity strength, we examined the test phase data for old scenes that were not endorsed as “recollect old,” and assessed whether there was a linear relationship between familiarity confidence and first saccade accuracy (Table 1). That is, first saccade accuracy was compared across the gradient of responses ranging from “sure new” to “sure old” (denoted by the regression line over the gray points in Figure 3a). First saccade accuracy did not differ across memory strength, β=.02, *t*(633)= 0.63, *p*=.54, indicating that it was not affected by familiarity, with substantial evidence for the null hypothesis (BF$_{10}$= .11).

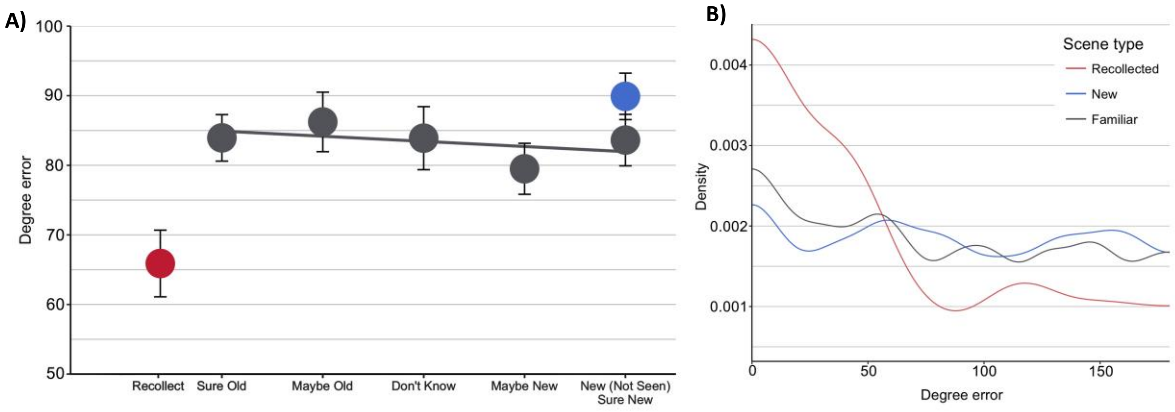


Figure 3. (A) Test phase average first saccade accuracy (the angular degree error between the first saccade’s direction from the center, and the direction of the target from the center; Table 2) by memory response. The “New (Not Seen)” (blue) scenes were correct rejections (i.e., “sure new” responses) of new scenes, which are included as a memory-less control; all other data points were old scenes, and these responses therefore include hits (recollect, sure old, and maybe old) and misses (don’t know, maybe new, sure new). Least-squares means controlling for participant and image are plotted, and the error bars represent the standard error of these estimated means from the model. The gray line represents the regression line across the familiarity-based responses used in the analysis of memory strength effects. (B) Density plot (i.e., smoothed histogram) showing the distribution of first saccade accuracy by memory type. Density corresponds to a smoothed estimate, using a Gaussian smoothing kernel, of scaled frequency[2] (see Appendix for a standard histogram).

---

[2] Figure 3b was produced from the histogram of first saccade accuracies, using smoothed density estimates to aid the eye in detecting trends. The curve was generated using a Gaussian smoothing kernel, and bin widths were set to 1/3 of the default for the kernel (as determined by the stat_density function within the ggplot2 package in R). The standard histogram, also using scaled frequency to facilitate comparison between categories with different numbers of trials, is plotted in Figure A.2.

To further characterize the relationship between first saccade accuracy and memory, we examined the density plot (i.e., smoothed histogram) of the distributions of first saccade accuracy by memory type (Figure 3b). The figure shows that recollected scenes were associated with a large proportion of very accurate trials, such that the first saccades in these trials were within approximately 60 degrees of the direct path to the target. For the recollected scenes, the trials that did not have these highly accurate first saccades appeared to have degree errors distributed roughly evenly across degree error values greater than 60 degrees, which may represent a distribution of guessing-based responses. That is, the distribution of degree error for recollected scenes appears visually analogous to a mixture model consisting of a highly precise distribution (i.e., <60 degrees), and a guessing distribution (Zhang & Luck, 2008). This suggests that recollection may improve average first saccade accuracy by causing a subset of first saccades to be highly accurate, rather than causing a diffuse, more incremental improvement in first saccade accuracy across all recollected trials. In contrast, the distributions for old scenes that were not recollected and new scenes were similar to each other and generally inaccurate, leading to distributions spread across the full range of degree error values. This suggests that for scenes that were not recollected, first saccades were not systematically related to the target location, and therefore did not appear to be guided by memory.

Scanpath efficiency. Figure 4a presents test phase scanpath efficiency, quantified as scanpath ratio, as a function of memory response and type of scene. The same memory contrasts and scenes that were used for the analyses of first saccade accuracy were used for the analyses of scanpath ratio as well (Table 1). Within scenes given “sure new” responses, scanpath ratio was significantly lower for old scenes than for new scenes, t(331)= -3.76, *p*<.001, *d*= -0.42—that is, scanpath ratio was lower for high-confidence misses than for high-confidence correct rejections. This indicates that unconscious memory for a scene led to a more efficient path to the target, even when participants were highly confident that they had not viewed the scene previously. In contrast, scanpath ratio was not improved by recollection compared to high-confidence familiarity, *t*(307.4)= -1.47, *p*=.14, *d*= -0.17, nor by familiarity strength overall, β= -.02, *t*(908)= -0.67, *p*=.51.

When scanpath ratio was recalculated from the second saccade onward to provide complete temporal separation from first saccade accuracy (i.e., excluding the first saccade), the same pattern of results was obtained, such that only unconscious memory improved efficiency (unconscious memory: *t*(375)= -2.28, *p*=.02, *d*= -0.24; recollection: *t*(361)= -0.91, *p*=.36, *d*= -0.10; familiarity: β= -.03, *t*(1022)= -0.98, *p*=.33). Furthermore, Bayesian analysis provided substantial evidence for the null effect of both conscious recollection (BF_{10}=.15) and familiarity (BF_{10}=.14), such that in each case, the numerical differences in scanpath ratio were more than six times more likely to be explained by chance than by recollection or familiarity. Thus, memory-related increases in scanpath efficiency were not observed for conscious recollection or memory strength, but rather were limited to unconscious memory.

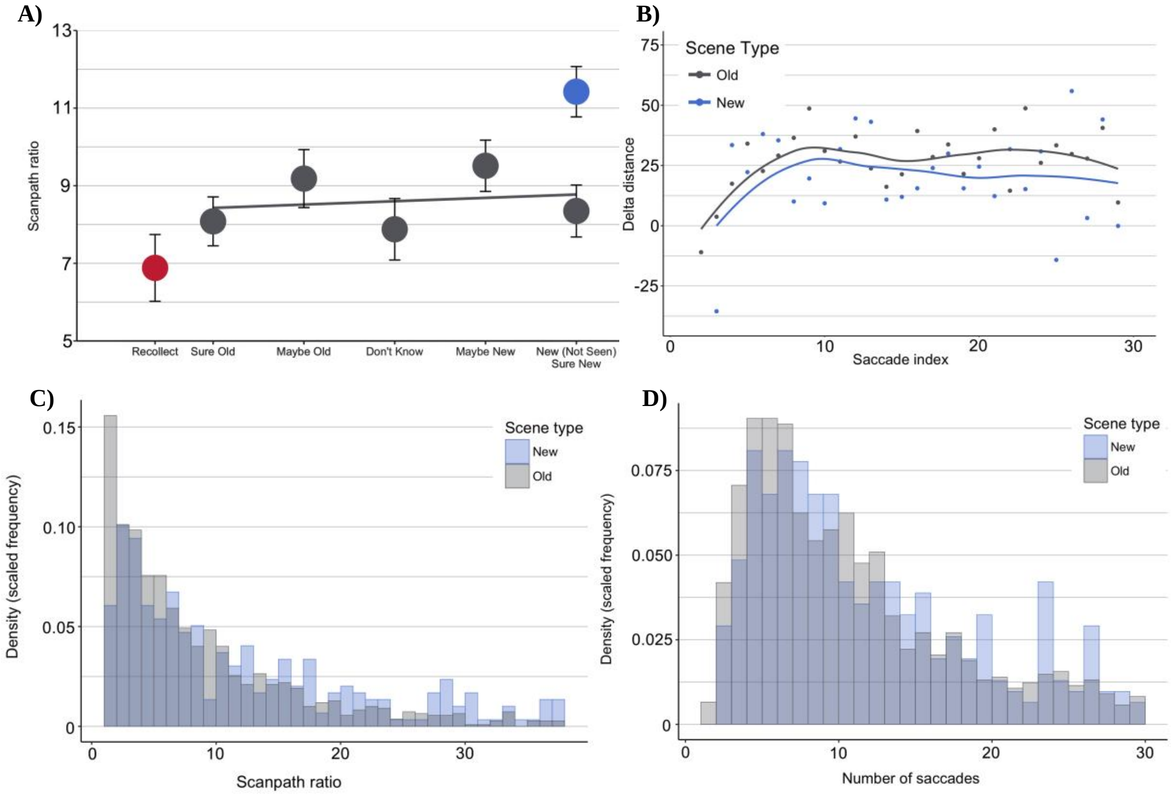


Figure 4. (A) Test phase average scanpath ratio (measured as the ratio of the observed scanpath to the shortest path to the target; Table 2) by memory response. The "New (Not Seen)" (blue) scenes were correct rejections (i.e., "sure new" responses) of new scenes, which are included as a memory-less control; all other data points were old scenes, and these responses therefore include hits (recollect, sure old, and maybe old) and misses (don't know, maybe new, sure new). Least-squares means controlling for participant and image are plotted, and the error bars represent the standard error of these estimated means from the model. The gray line represents the regression line across the familiarity-based responses used in the analysis of memory strength effects. (B) Delta distance, measured as the extent to which each saccade brought the eyes closer to the target, over the course of the trial (Table 2). The x-axis represents the saccade index, which is the ordinal saccade number in a trial. Each data point represents the average change in distance from the target (in pixels) for a given saccade index, with a separate data point for each condition (old versus new scenes). Locally weighted smoothing was used to produce the curve between data points in each condition.[3] (C) Histogram of scanpath ratio values for old and new scenes. (D) Histogram of the number of saccades made in a trial for old and new scenes. Scaled frequency values are presented in both (C) and (D) to facilitate comparisons between the old and new distributions, given that there were different numbers of old and new scenes. In each histogram, the bin width is one unit. Plots (B) – (D) include new scenes given a response of "sure new" (corresponding to the blue point in (A)), and all old scenes.

---

[3] In Figure 4b, the lines were generated using a loess smoothing (i.e., locally weighted smoothing) function in the ggplot2 package in R, which plots local regressions to aid the eye in seeing trends from scatterplots that may not necessarily be best captured by a standard linear regression.

Although analysis of scanpath ratio indicated that unconscious memory increased the overall efficiency of the eyes in a trial, the underlying mechanism is unclear: it is equally plausible that the improvement in scanpath ratio was simply driven by a reduction in the total number of saccades made in a trial, or by another mechanism such as an enhancement of the efficiency of each individual saccade. A secondary analysis aimed at testing this showed that the effect of unconscious memory on scanpath ratio was not solely driven by a reduction in the number of saccades needed to reach the target. Specifically, the number of saccades made in a trial was only marginally reduced by unconscious memory, $t(371)= -1.90$, $p=.05$ (Figure 4d). In contrast, when the number of saccades made in a trial was covaried in the model assessing the influence of unconscious memory on scanpath ratio, the effect of unconscious memory on scanpath ratio not only remained, but was in fact strengthened, $t(310)= -4.11$, $p<.0001$, $d=-.47$. Therefore, unconscious memory appeared to guide the eyes such that each saccade was more efficient, rather than simply reducing the number of saccades needed to reach the target.

Decomposing scanpath efficiency. The finding that scanpath ratio was significantly better for old scenes endorsed as “sure new” compared to new scenes endorsed as “sure new” indicates that unconscious influences of memory led to a more efficient scanpath. One interpretation of this effect is that the eyes are gradually guided toward the target throughout the search process by unconscious memory. However, another possibility is a two-state account whereby participants either have memory for the target location and rapidly move towards it, or they have no memory (conscious or unconscious) for the target location and effectively search for the target at random (e.g., Johnson et al., 2007; Smyth & Shanks, 2008)—and the existence of these two subsets of old trials would not be apparent when examining the average effects. An effect analogous to this two-state account was seen in the first saccade accuracy results above, such that recollection appeared to improve average first saccade accuracy by causing a subset of recollected trials to have highly precise first saccades, with the remaining trials appearing to be based on guesses (Figure 3b). If a similar two-state account were the case for scanpath efficiency, scanpath ratio for the majority of old scenes might be similar to that of the new scenes (i.e., there was no effect of memory for these scenes at all), but there may exist a subset of old scenes for which there was unconscious memory and a very highly efficient scanpath. For these latter trials, the eyes would have moved quickly to the target location, and therefore would be expected to elicit very few saccades and highly efficient scanpaths. To assess these possibilities, we considered two sets of analyses.

First, to quantify scanpath efficiency throughout the search process, we examined *delta distance*—the extent to which each saccade brought the eyes closer to the target—which allowed us to determine if memory led to consistently elevated efficiency throughout the course of search, for trials eliciting different numbers of saccades. Delta distance was calculated separately for each saccade in the test phase, and was defined as the change in distance from the target from the beginning to the end of the saccade (Table 2). As such, a larger delta distance value indicates that the saccade brought the eyes closer to the target. Given that the effects of memory on scanpath ratio were observed across all memory responses for old scenes, such that previous viewing of a scene improved scanpath ratio irrespective of memory response, all old scenes were included in the model of delta distance. If only a subset of trials involved direct guidance to the target whereas the rest proceeded as random search, as in the two-state account, one would expect delta distance to be greater for old scenes *only* at low saccade indexes (i.e., shorter trials).

Figure 4b shows that delta distance was greater for old scenes than for new scenes, and this effect was observed across both low and high saccade indexes. That is, there was a significant effect of old/new status of a scene, *t*(16400)= 3.15, *p*=.002, *d*= 0.05, and no interaction between old/new status and saccade index, *t*(16461)= -0.37, *p*=.71. This indicates that memory enhanced scanpath efficiency similarly for both early and late saccades, and for trials ranging from few to many saccades. Similar results were observed when we restricted the analysis to scenes that were endorsed as "sure new," such that there was a significant effect of old/new status, *t*(6527)=2.72, *p*=.007, *d*= 0.07, and no old/new by saccade index interaction, *t*(6517)= -0.50, *p*=.62. Lastly, when analysis was restricted to old scenes, there was no effect of any memory response on delta distance, β= -.0004, *t*(12370)= 0.05, *p*=.96, suggesting that improvements in delta distance were uniquely related to unconscious memory, without an influence of familiarity or conscious recollection. Therefore, examination of delta distance suggests that unconscious memory gradually led the eyes to the target throughout the course of search. (For a related analysis that controls for the total number of saccades made in a trial, see Appendix.)

Second, we examined the histograms of scanpath ratio and number of saccades elicited by old and new scenes in the test phase, to assess whether there was evidence for two different underlying distributions of scanpaths for the old scenes. That is, according to the two-state account, one might expect to see a majority of trials with high scanpath ratios and many saccades, reflecting random search equivalent to the search process observed in new scenes—along with a subset of old trials with scanpath ratios near ceiling (i.e., around 1) and few saccades, reflecting that the eyes were guided directly to the target. This two-state account would be corroborated by distributions that are similar to those seen in first saccade accuracy (Figure

3b), such that a subset of old trials are near ceiling, and the remaining distribution is an evenly reduced version of the new trial distribution. That is, the subset of near-ceiling trials would have appeared to be evenly sampled from the remainder of the distribution rather than resulting from an overall shift of the distribution towards higher performance. However, instead of this kind of mixture distribution, both Figure 4c and Figure 4d suggest that the distribution of old trials is simply shifted left towards lower scanpath ratios and fewer saccades, respectively, than new trials. As such, the distributions do not suggest that there was a mixture of trials with memory leading the eyes directly to the target and trials that are indistinguishable from new trials. Instead, the distributions suggest that there was an increase in search efficiency affecting the majority of trials, in contrast to the effects of recollection on first saccade accuracy.

Replication. All of the test phase analyses above were conducted using the scenes that were presented once in the learning phase, as we expected the scenes that were presented three times to be at ceiling levels of recognition. However, we also examined performance on these latter scenes for the sake of completeness, and the analyses largely replicated the effects we observed in the scenes presented once. All significant and null effects replicated for degree error of first saccade (recollection effect: $t(431)= -3.01$, $p=.003$, $d= -0.29$), scanpath ratio (unconscious effect: $t(241)= -2.48$, $p=.013$, $d= -0.32$), and delta distance (old/new effect: $t(4213) = 6.86$, $p<.0001$, $d= 0.21$), except for the null interaction between saccade index and old/new status in delta distance, $\beta = -.03$, $t(6174)= -3.62$, $p<.001$. This interaction indicates that the magnitude of the effect of old/new status on delta distance changed over the course of the trial or across trials of different lengths, which could suggest that the influence of unconscious memory on eye movement guidance was restricted to a particular point in the search process or to trials consisting of a certain number of saccades. However, examination of the data revealed that this significant interaction was instead due to a transient reduction in the effect of old/new status on delta distance around the 21^{st}-25^{th} saccade (likely due to having few trials of that length for scenes presented three times), whereas old/new status showed a consistent effect throughout the rest of the saccades. Therefore, analyses using the scenes presented three times provided an internal replication of the main results outlined above.

Discussion

In the present study, we examined eye movements made during a visual search task and used a subjective report procedure that isolated the contributions of conscious recollection, unconscious memory, and familiarity strength. The present results demonstrate that eye movements made while viewing real-world scenes can be influenced by both conscious and unconscious memory. Whereas conscious recollection for a scene uniquely improved the accuracy of the first eye movement in a search task, unconscious memory uniquely improved participants' search efficiency and gradually guided the eyes towards the target over the course of a trial. Examination of the distributions underlying these effects suggested that conscious recollection may lead to very highly accurate first saccades on a subset of trials, whereas unconscious memory may instead lead to a general and diffuse improvement in the efficiency of each saccade on the majority of trials. Furthermore, Bayesian analyses indicated that these memory effects on eye movements may be independent, such that conscious memory did not influence scanpath efficiency, and unconscious memory did not influence first saccade accuracy—and indicated that familiarity strength did not influence either of these patterns of eye movements. These results build upon previous work by introducing a novel paradigm that allows for the separation of conscious and unconscious memory from intermediate strength-based memory that could bias these measures, and, through eyetracking, allows for the decomposition of previously reported reaction time effects.

By demonstrating that multiple types of memory can influence attention simultaneously, the present results may be able to explain past conflicting findings on whether contextual cueing effects are due to conscious or unconscious memory. That is, the present finding of multiple forms of memory contributing to contextual cueing suggests that the contradictory conclusions reached in prior studies may be due to the dichotomous old/new memory assessments used, and to differences in memorability of the stimuli used. Specifically, many of the studies reporting that contextual cueing is driven only by unconscious memory, due to the absence of conscious memory for the stimuli, used abstract arrays of letters as the background context (e.g., Chun, 2000; Chun & Jiang, 1998; Goujon et al., 2015). Abstract arrays generally do not elicit levels of conscious memory that are detectable in the dichotomous old/new memory tests that are used in these studies, likely due to very high similarity between stimuli and a lack of semantic information. On the other hand, those studies attributing contextual cueing to conscious memory, due to very high recognition accuracy on dichotomous memory tests, have typically used scenes as stimuli (e.g., Brockmole & Henderson, 2006a, 2006b)—for which people have exceptionally good memory (Konkle, Brady, Alvarez, & Oliva, 2010; Standing, 1973). Given that contextual cueing effects are stronger and appear sooner in scenes than in arrays, combined with the present results, it seems likely that contextual cueing in scenes is driven by both conscious and unconscious memory, whereas contextual cueing in abstract arrays of letters may be driven primarily by unconscious memory.

The finding that conscious recollection influenced the first saccade whereas unconscious memory influenced performance throughout the course of search might seem surprising, given that conscious influences of memory are generally thought to be slower than unconscious influences (Schacter, Wagner, & Buckner, 2000; Yonelinas, 2002). We suspect that one reason that the conscious memory effects occurred so early in search is because participants saw a preview of and made explicit memory judgments about each test phase scene prior to searching for the target, and as such, eye movements related to conscious recollection may have been planned prior to the onset of the target-containing version of the scene. However, while the explicit recognition response likely did impact the ability of recollection to influence the first eye movement, the same effects on the first saccade were already apparent in the learning phase, where there was no scene preview or memory judgment—suggesting that these early eye movements are observed even without a preview or explicit retrieval demands. We interpret the results as indicating that conscious recollection can impact eye movements relatively quickly after stimulus onset, even without a preview.

The current results do not speak directly to questions regarding the neural substrates of these effects, but we believe that combining the present subjective report methods with patient studies or neuroimaging methods may be useful in addressing these questions. For example, there is conflicting evidence regarding whether the hippocampus or regions in the surrounding medial temporal lobe are involved in supporting memory-related patterns of eye movements (Chun & Phelps, 1999; Giesbrecht, Sy, & Guerin, 2013; Manns & Squire, 2001; Preston & Gabrieli, 2008), and the present paradigm may prove useful in informing this debate when combined with measures of neural substrates. Furthermore, a growing body of memory research has indicated that recollection is dependent on the hippocampus (Eichenbaum et al., 2007), and that it is associated with relatively high-resolution information (Yonelinas, 2013). Specifically, the hippocampus has been shown to provide high-resolution spatial information about previous events, whereas cortical regions outside the hippocampus can support less precise spatial information such as the general quadrant in which objects were previously encountered (Koen, Borders, Petzold, & Yonelinas, 2017; Kolarik et al., 2016). The fact that recollection in the current study was associated with highly accurate first saccades towards the target location is consistent with the proposed role of the hippocampus in supporting high-precision memory responses.

Whereas the deployment of highly precise memory representations supported by the hippocampus is a feasible mechanism through which recollection may influence first saccade accuracy, the potential mechanisms underlying the influence of unconscious memory on scanpath efficiency are less easily identified. There is evidence, however, for relational memory that can be expressed through eye movements in the absence of awareness for that memory (Hannula et al., 2010; Ryan et al., 2000), and it has been proposed that relational memory is instrumental in binding targets to contexts in support of contextual cueing (Chun & Phelps, 1999). It is therefore possible that the present findings of guidance by unconscious memory may reflect relational memory, which could use contextual scene information to gradually guide the eyes towards the region containing the target. Another possibility is that unconscious memory may improve efficiency throughout the search process by reducing the extent to which saccades are made to incorrect regions. That is, perhaps increased fluency of perceptual processing, resulting from prior exposure (Schacter et al., 2000; Voss & Paller, 2008), reduces the amount of visual information needed to reject incorrect regions. Studies aimed at identifying the mechanisms underlying the present effects would be useful in determining how unconscious memory influences attention.

In addition to furthering our understanding of how experience guides eye movements, these findings pave the way for a variety of practical applications. For example, extending this work to real-world tasks could allow us to understand and enhance the influence of experience on radiologists’ ability to detect cancers in x-ray images, or TSA agents’ ability to find weapons in luggage (Wolfe, Brunelli, Rubinstein, & Horowitz, 2013; Wolfe, 2016): if these results extend to classes of images in addition to repeated instances of an image, perhaps eye movement measures could detect unconscious forms of memory-informed search that are not available to conscious report. Moreover, to the extent that eye movements can be used to index different forms of memory, these measures may provide an efficient means to quickly detect different types of memory impairments such as those observed in aging and Alzheimer’s disease. Future research is needed to determine whether the presently identified patterns of eye movements may generalize to other conditions, such that they may serve as indicators of conscious and unconscious memory influences.

References

Althoff, R. R., & Cohen, N. J. (1999). Eye-Movement-Based Memory Effect: A Reprocessing Effect in Face Perception. *Journal of Experimental Psychology: Learning Memory and Cognition*, *25*(4), 997–1010. http://doi.org/10.1037/0278-7393.25.4.997

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, *67*(1), 1–48. http://doi.org/10.18637/jss.v067.i01

Brockmole, J. R., & Henderson, J. (2006a). Recognition and attention guidance during contextual cueing in real-world scenes: Evidence from eye movements. *Quarterly Journal of Experimental Psychology*, *59*(7), 1177–1187. http://doi.org/10.1080/17470210600665996

Brockmole, J. R., & Henderson, J. M. (2006b). Using real-world scenes as contextual cues for search. *Visual Cognition*, *13*(1), 99–108. http://doi.org/10.1080/13506280500165188

Castelhano, M. S., & Henderson, J. M. (2007). Initial Scene Representations Facilitate Eye Movement Guidance in Visual Search. http://doi.org/10.1037/0096-1523.33.4.753

Chun, M. M. (2000, May). Contextual cueing of visual attention. *Trends in Cognitive Sciences*. http://doi.org/10.1016/S1364-6613(00)01476-5

Chun, M. M., & Jiang, Y. (1998). Contextual Cueing: Implicit Learning and Memory of Visual Context Guides Spatial Attention. *Cognitive Psychology*, *36*, 28–71. Retrieved from http://resolver.scholarsportal.info/resolve/00100285/v36i0001/28_ccilamovcgsa.xml

Chun, M. M., & Phelps, E. A. (1999). Memory deficits for implicit contextual information in amnesic subjects with hippocampal damage. *Nature Neuroscience*, *2*, 844–847. http://doi.org/10.1038/12222

Donaldson, W. (1996). *The role of decision processes in remembering and knowing*. *Memory &*

Cognition (Vol. 24).

Eichenbaum, H., Yonelinas, A. P., & Ranganath, C. (2007). The medial temporal lobe and recognition memory. *Annual Review of Neuroscience, 30*, 123–152. http://doi.org/10.1146/annurev.neuro.30.051606.094328

Giesbrecht, B., Sy, J. L., & Guerin, S. A. (2013). Both memory and attention systems contribute to visual search for targets cued by implicitly learned context. *Vision Research*, *85*, 80–89. http://doi.org/10.1016/j.visres.2012.10.006

Goujon, A., Didierjean, A., & Marmèche, E. (2007). Contextual cueing based on specific and categorical properties of the environment. *Visual Cognition*, *15*(3), 257–275. http://doi.org/10.1080/13506280600677744

Goujon, A., Didierjean, A., & Thorpe, S. (2015). Investigating implicit statistical learning mechanisms through contextual cueing. *Trends in Cognitive Sciences*, *19*(9), 524–533. http://doi.org/10.1016/j.tics.2015.07.009

Hannula, D. E. (2010). Worth a glance: using eye movements to investigate the cognitive neuroscience of memory. *Frontiers in Human Neuroscience, 4,* 166. http://doi.org/10.3389/fnhum.2010.00166

Hannula, D. E., Baym, C. L., Warren, D. E., & Cohen, N. J. (2012). The eyes know: Eye movements as a veridical index of memory. *Psychological Science*, *23*(3), 278–287. http://doi.org/10.1177/0956797611429799

Hannula, D. E., Ranganath, C., Ramsay, I. S., Solomon, M., Yoon, J., Niendam, T. A., … Ragland, J. D. (2010). Use of eye movement monitoring to examine item and relational memory in schizophrenia. *Biological Psychiatry, 68*(7), 610–616. http://doi.org/10.1016/j.biopsych.2010.06.001

Henderson, J. M., & Hollingworth, A. (2003). Eye movements and visual memory: Detecting changes to saccade targets in scenes. *Perception and Psychophysics*, *65*(1), 58–71. http://doi.org/10.3758/BF03194783

Hollingworth, A., Williams, C. C., & Henderson, J. M. (2001). To see and remember: Visually specific information is retained in memory from previously attended objects in natural scenes. *Psychonomic Bulletin and Review*, *8*(4), 761–768. http://doi.org/10.3758/BF03196215

Jeffreys, H. (1961). *Theory of probability* (3rd ed.). Oxford, England: Clarendon.

Johnson, J. S., Woodman, G. F., Braun, E. L., & Luck, S. J. (2007). Implicit memory influences the allocation of attention in visual cortex. *Psychonomic Bulletin & Review*, *14*(5), 834–839. http://doi.org/10.3758/BF03194108

Kafkas, A., & Montaldi, D. (2011). Recognition memory strength is predicted by pupillary responses at encoding while fixation patterns distinguish recollection from familiarity. *The Quarterly Journal of Experimental Psychology*, *64*(10), 1971–1989. http://doi.org/10.1080/17470218.2011.588335

Kafkas, A., & Montaldi, D. (2012). Familiarity and recollection produce distinct eye movement, pupil and medial temporal lobe responses when memory strength is matched. *Neuropsychologia*, *50*(13), 3080–3093. http://doi.org/10.1016/j.neuropsychologia.2012.08.001

Koen, J. D., Borders, A. A., Petzold, M. T., & Yonelinas, A. P. (2017). Visual short-term memory for high resolution associations is impaired in patients with medial temporal lobe damage. *Hippocampus*, *27*, 184–193. http://doi.org/10.1002/hipo.22682

Kolarik, B. S., Shahlaie, K., Hassan, A., Borders, A. A., Kaufman, K. C., Gurkoff, G., …

Ekstrom, A. D. (2016). Impairments in precision, rather than spatial strategy, characterize performance on the virtual Morris Water Maze: A case study. *Neuropsychologia*, *80*, 90–101. http://doi.org/10.1016/j.neuropsychologia.2015.11.013

Konkle, T., Brady, T. F., Alvarez, G. A., & Oliva, A. (2010). Scene memory is more detailed than you think: The role of categories in visual long-term memory. *Psychological Science*, *21*(11), 1551–1556. http://doi.org/10.1177/0956797610385359

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest Package: Tests in Linear Mixed Effects Models. *Journal of Statistical Software*, *82*(13), 1–19. http://doi.org/10.18637/jss.v082.i13

Lenth, R. V. (2016). Least-Squares Means: The R Package lsmeans. *Journal of Statistical Software*, *69*(1). http://doi.org/10.18637/jss.v069.i01

Manns, J. R., & Squire, L. R. (2001). Perceptual learning, awareness, and the hippocampus. *Hippocampus*, *11*(6), 776–782. http://doi.org/10.1002/hipo.1093

Peterson, M. S., & Kramer, A. F. (2001). Attentional guidance of the eyes by contextual information and abrupt onsets. *Perception & Psychophysics, 63*(7), 1239–1249. http://doi.org/10.3758/BF03194537

Preston, A. R., & Gabrieli, J. D. E. (2008). Dissociation between explicit memory and configural memory in the human medial temporal lobe. *Cerebral Cortex*, *18*, 2192–2207. http://doi.org/10.1093/cercor/bhm245

Rosenthal, R., & Rosnow, R. L. (1991). *Essentials of Behavioral Research: Methods and Data Analysis*. *McGraw-Hill Series in Psychology*. Retrieved from https://lggy89ten01.storage.googleapis.com/MDA3MzUzMTk2MA==01.pdf

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. http://doi.org/10.1111/1467-9280.00288

Ryan, J. D., Hannula, D. E., & Cohen, N. J. (2007). The obligatory effects of memory on eye movements. *Memory*, *15*(5), 508–525. http://doi.org/10.1080/09658210701391022

Schacter, D. L. ., Wagner, A. D. ., & Buckner, R. L. (2000). Memory Systems of 1999. *The Oxford Handbook of Memory*, 627–643.

Schlagbauer, B., Muller, H. J., Zehetleitner, M., & Geyer, T. (2012). Awareness in contextual cueing of visual search as measured with concurrent access- and phenomenal-consciousness tasks. *Journal of Vision*, *12*(11), 1–12. http://doi.org/10.1167/12.11.25

Sharot, T., Davidson, M. L., Carson, M. M., & Phelps, E. A. (2008). Eye movements predict recollective experience. *PloS ONE*, *3*(8). http://doi.org/10.1371/journal.pone.0002884

Smith, C. N., Hopkins, R. O., & Squire, L. R. (2006). Experience-dependent eye movements, awareness, and hippocampus-dependent memory. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *26*(44), 11304–12. http://doi.org/10.1523/JNEUROSCI.3071-06.2006

Smith, C. N., & Squire, L. R. (2008). Experience-Dependent Eye Movements Reflect Hippocampus-Dependent (Aware) Memory. *Journal of Neuroscience*, *28*(48), 12825–12833. http://doi.org/10.1523/JNEUROSCI.4542-08.2008

Smith, C. N., & Squire, L. R. (2017). When eye movements express memory for old and new scenes in the absence of awareness and independent of hippocampus. *Learning & Memory*, *24*(2), 95–103. http://doi.org/10.1101/lm.043851.116

Smyth, A. C., & Shanks, D. R. (2008). Awareness in contextual cuing with extended and concurrent explicit tests. *Memory and Cognition*, *36*(2), 403–415. http://doi.org/10.3758/MC.36.2.403

Standing, L. (1973). Learning 10,000 pictures. *Quarterly Journal of Experimental Psychology*, *25*(2), 207–222. http://doi.org/10.1080/14640747308400340

Voss, J. L., & Paller, K. A. (2008). Brain substrates of implicit and explicit memory: The importance of concurrently acquired neural signals of both memory types. *Neuropsychologia*, *46*(13), 3021–3029. http://doi.org/10.1016/j.neuropsychologia.2008.07.010

Wolfe, J. M. (2016). Use-inspired basic research in medical image perception. *Cognitive Research: Principles and Implications*, 1–17. http://doi.org/10.1186/s41235-016-0019-2

Wolfe, J. M., Brunelli, D. N., Rubinstein, J., & Horowitz, T. S. (2013). Prevalence effects in newly trained airport checkpoint screeners: trained observers miss rare targets, too. *Journal of Vision*, *13*(3), 1–9. http://doi.org/10.1167/13.3.33

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

Yonelinas, A. P. (2013). The hippocampus supports high-resolution binding in the service of perception, working memory and long-term memory. *Behavioural Brain Research*, *254*, 34–44. http://doi.org/10.1016/j.bbr.2013.05.030

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

Appendix

Statistical Analyses

Correlations between search speed reaction time (RT) and eye movement measures. To further verify that the eye movement measures used were reflective of search performance, and therefore contextual cueing performance, we determined whether the eye movement measures were related to RT. An analysis of all of the trials in which the target was identified, throughout both the learning and test phases, showed a significant correlation between scanpath ratio and RT (r=.75, p<.0001). The correlation between first saccade accuracy and RT was small, but also significant (r=.12, p<.0001). Despite the weaker relationship between first saccade accuracy and RT (likely driven by the much smaller proportion of total trial time indexed by first saccade accuracy), RT appears to be related to first saccade accuracy independently from improvements in scanpath ratio, as first saccade accuracy remains significant (p<.0001) in predicting RT even when scanpath ratio is controlled for.

Influence of encoding time. Due to the possibility of confounds stemming from the expected relation between higher initial viewing time (i.e., RT on first presentation) and stronger memory, we repeated all linear mixed effects models (Table 1) assessing the relation between memory and eye movement patterns with the additional covariate of RT for the first presentation of the scene. This was not found to alter the pattern of results.

Scanpath ratio. Scanpath ratio values were skewed; to ensure that skew did not significantly impact the observed effects, we re-ran the analyses with log-transformed scanpath ratio values, and found that the unconscious effect held (p<.001) as did the null effects of familiarity and recollection. The other variables of interest were not skewed.

The first calculation of scanpath ratio presented in the main text included data from the first saccade. To provide statistical separation of scanpath ratio from first saccade accuracy, models of the effect of memory on this measure of scanpath ratio included a covariate of first saccade accuracy. That is, Figure 4a and the statistics reported in the main text assessed the influence of memory on scanpath ratio for a given degree of first saccade accuracy. The null effect of recollection on scanpath ratio was similar regardless of whether or not this covariate was included, however (with covariate, p=.14, without covariate: p=.12). Furthermore, inclusion of the first saccade accuracy covariate did not alter the pattern of results with respect to familiarity nor unconscious memory. This covariate was not included in the models assessing scanpath ratio that was calculated from the second saccade onward.

Decomposition of scanpath efficiency: additional analysis. Delta distance was computed to determine whether scanpath efficiency was elevated throughout the course of the search process, and for trials consisting of different numbers of saccades. In order to have a sufficient number of trials in each saccade index bin, trials consisting of 30 or fewer saccades were used in this analysis. However, because different trials required different numbers of saccades to locate the target, plotting all of these trials together may obscure effects that vary depending on trial length. To observe the trends in unconscious memory guidance throughout trials consisting of a similar number of saccades, we plotted *absolute* distance from the target, over the course of saccades, for trials in which the target was found within 6-8 saccades (Figure A.1). (The mode number of saccades in a trial was 7; Figure 4d) This plot corroborates the conclusion that unconscious memory gradually guided the eyes towards the target over the course of the trial, as the eyes appear to be consistently closer to the target in old scenes.

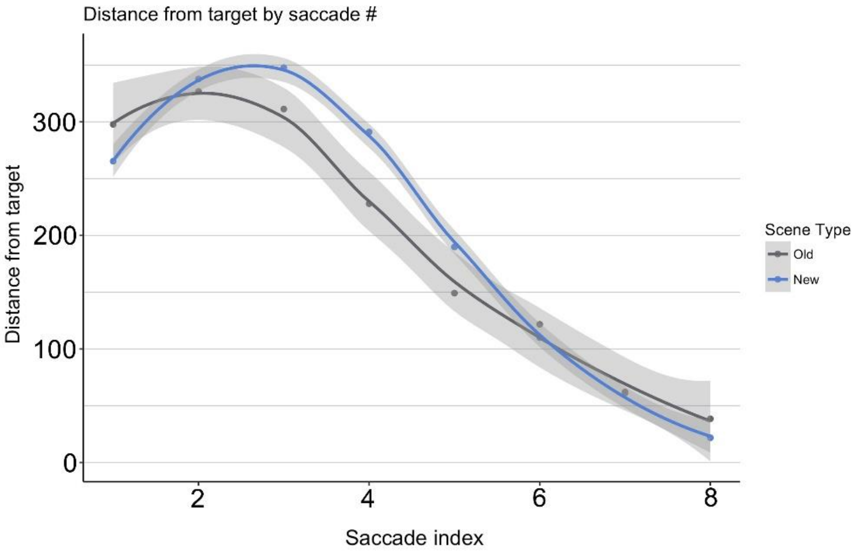


Figure A.1. Absolute distance from the target over the course of the trial, for trials in which the target was found within 6-8 saccades. All scenes were endorsed as “sure new,” and are sorted by whether or not they appeared in the learning phase. Shaded regions represent the standard error of the mean. The same smoothing method was used as that in Figure 4b.

Plotting. Figures 2, 3a and 4a were plotted using least-squares means obtained using the linear mixed effects models. Thus, the plots control for participant and image as in the models reported in the text—but the trends are similar when the raw data is plotted instead. Error bars in these figures were calculated using the standard error values generated by the lsmeans package in R (Lenth, 2016).

Figure 3b presents a density plot of first saccade accuracies, but a standard version of the histogram is presented below (Figure A.2). Proportions (i.e., scaled frequency) are used to facilitate comparison between distributions containing different numbers of trials.

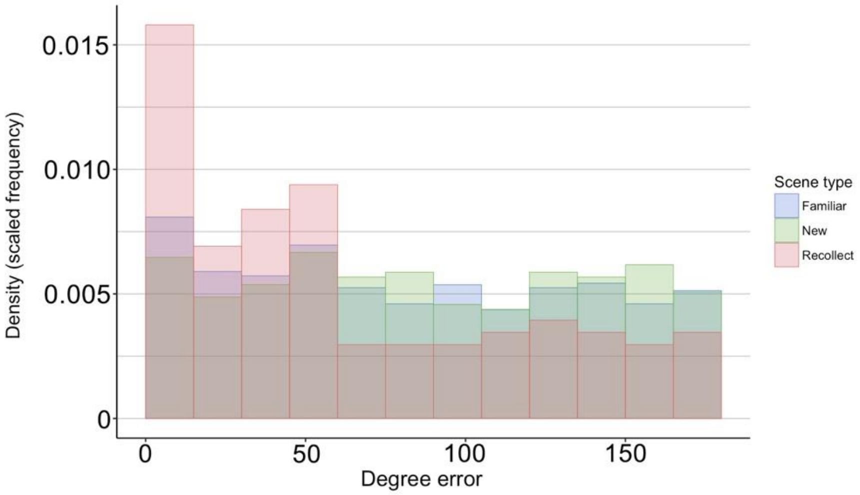


Figure A.2. Histogram of first saccade accuracy in the test phase, sorted by scene type. "New" scenes include all new scenes; "Familiar" scenes include old scenes presented once that were not recollected, and "Recollect" scenes include old scenes presented once that were recollected.

Degrees of freedom. Because the degrees of freedom were derived from the linear mixed effects models using the Satterthwaite approximation, the degrees of freedom were based on the number of trials included in an analysis rather than the number of participants included in the experiment. Therefore, a high value for degrees of freedom reflects the fact that a large number of data points were included in a given model. Degrees of freedom obtained using this method often contain decimals, but they were rounded to the nearest integer in the manuscript.

Chapter 2

The spatial distribution of attention predicts familiarity strength during encoding and retrieval

Michelle M. Ramey[a,b,c*], John M. Henderson[a,c] Andrew P. Yonelinas[a,b]

[a] Department of Psychology, University of California, Davis, CA, USA.

[b] Center for Neuroscience, University of California, Davis, CA, USA.

[c] Center for Mind and Brain, University of California, Davis, CA, USA.

Published in *Journal of Experimental Psychology: General* (2020)

Author Note:

Funding: This work was supported by the National Eye Institute of the National Institutes of Health under Award Numbers R01EY025999 and R01EY027792. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Abstract

The memories we form are determined by what we attend to, and conversely, what we attend to is influenced by our memory for past experiences. Although we know that shifts of attention via eye movements are related to memory during encoding and retrieval, the role of specific memory processes in this relationship is unclear. There is evidence that attention may be especially important for some forms of memory (i.e., conscious recollection), and less so for others (i.e., familiarity-based recognition and unconscious influences of memory), but results are conflicting with respect to both the memory processes and eye movement patterns involved. To address this, we used a confidence-based method of isolating eye movement indices of spatial attention that are related to different memory processes (i.e., recollection, familiarity strength, and unconscious memory) during encoding and retrieval of real-world scenes. We also developed a new method of measuring the dispersion of eye movements, which proved to be more sensitive to memory processing than previously used measures. Specifically, in two studies, we found that familiarity strength—that is, changes in subjective reports of memory confidence—increased with i) more dispersed patterns of viewing during encoding, ii) less dispersed viewing during retrieval, and iii) greater overlap in regions viewed between encoding and retrieval (i.e., resampling). Recollection was also related to these eye movements in a similar manner, though the associations with recollection were less consistent across experiments. Furthermore, we found no evidence for effects related to unconscious influences of memory. These findings indicate that attentional processes during viewing may not preferentially relate to recollection, and that the spatial distribution of eye movements is directly related to familiarity-based memory during encoding and retrieval.

Keywords: Recognition; implicit memory; recollection and familiarity; eye movements; memory

The spatial distribution of attention predicts familiarity strength during encoding and retrieval

How we view the world is influenced by our memory: for example, we know approximately where to look when searching for objects based on past experiences (Torralba, Oliva, Castelhano, & Henderson, 2006; Võ & Wolfe, 2015). Conversely, our ability to learn about new visual information depends on how we look at that information: where we direct our attention largely determines what we encode into memory (e.g., Henderson & Hollingworth, 1999, 2003; Hollingworth, 2006). This bidirectional relationship between memory and visual attention has long been supported by a wealth of research in both memory and attention (Hannula, 2010; Henderson, 2003; Meister & Buffalo, 2016). However, recent evidence indicates that understanding the relationship between memory and visual attention may require consideration of the type of memory involved, because different memory processes may be related to attention in different ways. For example, memory for prior experiences can be based on a variety of underlying memory processes, such as conscious *recollection* for details of an experience, assessments of stimulus *familiarity*, as well as *unconscious* forms of memory that occur without awareness (Roediger, 1990; Schacter, Chiu, & Ochsner, 1993; Yonelinas, 2001, 2002). A number of studies suggest that conscious recollection may have a uniquely strong relationship with visual attention (Holm & Mantyla, 2007; Kafkas & Montaldi, 2012; Mantyla & Holm, 2006; Sharot, Davidson, Carson, & Phelps, 2008), but other studies indicate that attentional mechanisms may be related to familiarity as well (Kafkas & Montaldi, 2011), and still others suggest that many experience-related changes in attention may be due to unconscious memory (Hannula, 2010; Hannula & Ranganath, 2009; Ryan, Althoff, Whitlow, & Cohen, 2000). In addition, there is emerging evidence that these memory processes may be differentially related to distinct aspects of spatial attention as indexed by eye movements (Ramey, Yonelinas,

& Henderson, 2019), but very little is known about the spatial attentional mechanisms involved in successful memory encoding and retrieval, even irrespective of memory processes. Resolving these issues is important both for accurately characterizing the functional nature of different memory processes, and for understanding the manner in which attentional processes influence—and are influenced by—memory for past experiences.

There are several ways in which visual attention, as indexed by eye movements, can influence or be influenced by memory during encoding and retrieval. For example, during encoding, attention can determine which image components are ultimately stored in memory (Henderson & Hollingworth, 1999, 2003; Hollingworth, 2006). In addition, memory for past encounters with an image may impact how we subsequently view that image (Althoff & Cohen, 1999; Ryan et al., 2000; Smith, Hopkins, & Squire, 2006; Smith & Squire, 2008, 2017), and conversely, preliminary evidence suggests that during memory retrieval, attention may influence memory by determining which areas of a scene are utilized as retrieval cues (Foulsham & Kingstone, 2013; Valuch, Becker, & Ansorge, 2013). However, it is not clear how the distribution of attention across scenes is related to different memory processes.

Examinations of eye movements during memory encoding and retrieval have thus far largely involved measures indexing general oculomotor behavior (e.g., the *number of fixations* made during viewing, and the average *saccade amplitude* or distance between fixations) rather than the distribution of visual attention per se. Nonetheless, these studies have provided useful clues as to how visual attention may relate to memory. For example, an increased number of fixations during encoding has been found to predict better overall memory performance on a subsequent memory test (Kafkas & Montaldi, 2011; Loftus, 1972; Sharot et al., 2008), suggesting that sampling more information, independent of encoding duration, may improve memory. However, studies aimed at determining how the number of fixations and average saccade amplitude during encoding predict subsequent recollection or familiarity have not yet yielded a clear conclusion. Specifically, one study found that the number of fixations made while viewing images was related to both recollection and familiarity strength (Kafkas & Montaldi, 2011), but a similar study found that the number of fixations was not specifically related to either memory process (Sharot et al., 2008). Both studies concluded that recollection was related to decreased saccade amplitude during encoding, which suggests that saccade amplitude may uniquely predict subsequent recollection.

At time of retrieval, previously studied images generally elicit fewer fixations than new images (Althoff & Cohen, 1999; Ryan et al., 2000; Smith et al., 2006; Smith & Squire, 2008, 2017), suggesting that memory leads to more efficient processing of repeated images. However, it is not yet clear how this effect relates to different memory processes. One study found that, compared to familiarity, recollection was related to decreased saccade amplitude and fewer fixations (Sharot et al., 2008), whereas another found that recollection was related to increased saccade amplitude and more fixations (Kafkas & Montaldi, 2012). In addition to conflicting results regarding how eye movements during retrieval may be related to recollection and familiarity, there is debate surrounding whether experience-driven changes in attention are instead a result of unconscious memory, rather than any form of conscious memory (Ryan et al., 2000; Smith et al., 2006; Smith & Squire, 2008, 2017). In a study directly examining how recollection, familiarity, and unconscious memory may drive experience-related changes in attention during an implicit visual search task, we found that recollection (i.e., conscious memory) and unconscious memory outside of awareness—but not familiarity—were each related to distinct spatial patterns of eye movements (Ramey et al., 2019). Specifically, we found that recollection was uniquely associated with increased accuracy of the first eye movement in a trial in terms of heading towards the learned location of a search target, and that unconscious memory was uniquely associated with an improvement in search efficiency throughout the trial. These results indicate that both recollection and unconscious memory may relate to distinct patterns of spatial attentional allocation, but it remains to be seen whether analogous findings apply during explicit encoding and retrieval.

Finally, investigations of the extent to which image regions are revisited between successive viewings have provided a rare, direct examination of spatial allocation of attention during encoding and retrieval. Overall memory tends to be better when participants resample scene regions at retrieval that they had initially visited at encoding (Foulsham & Kingstone, 2013; Wynn et al., 2016). Furthermore, this effect appears to be bidirectional, because forcing participants to resample more regions improves memory for a scene in some cases, and cueing memory prior to viewing increases the degree of resampling upon viewing (Foulsham & Kingstone, 2013; Holm & Mantyla, 2007). Importantly, there is evidence that this effect may be recollection-related, such that increased resampling has been found to uniquely relate to recollection, but familiarity strength has not yet been examined (Holm & Mantyla, 2007; Mantyla & Holm, 2006).

The notion that recollection, rather than familiarity, is particularly related to attentional allocation may seem intuitive given that recollection involves memory for specific details of an experience (Yonelinas, 2002), and eye movements allow for the extraction of detailed visual information (Henderson, 2003). Familiarity-based recognition, on the other hand, is posited to involve a sense of global similarity (Yonelinas, 2002), and the overall gist of visual stimuli may be ascertained without the need for eye movements (Henderson, 2003). Despite this, however, previous results using general oculomotor measures are inconsistent with respect to the directions of the associations between different eye movement patterns and recollection, and to what extent familiarity strength is involved (Kafkas & Montaldi, 2011, 2012; Sharot et al., 2008). Thus, the existing literature shows that visual attention is related to memory at both encoding and retrieval, but it is not yet clear how it is related to specific memory processes.

One possible reason for the conflicting findings is that memory strength has not been taken into account in most studies, which could cause familiarity strength effects to appear as recollection effects (Montaldi, Spencer, Roberts, & Mayes, 2006; Yonelinas, Otten, Shaw, & Rugg, 2005). That is, while recollection-based responses most often involve high-confidence recognition, familiarity-based responses vary widely in recognition confidence (Yonelinas et al., 2005). Therefore, when a dichotomous measure is used such that recollection is compared to all levels of familiarity confidence, as has most often been done in eye movement investigations of memory processes, observed relationships between attention and memory could be due to associations with overall memory strength instead of recollection. In fact, in one study that did assess memory strength, the reported relationship between the number of fixations and recollection did not hold when recollected stimuli were compared to stimuli with only high-strength familiarity, as opposed to all familiar stimuli (Kafkas & Montaldi, 2012). This result suggests that some effects interpreted to be recollection-specific may actually have been a result of the high memory strength that recollection entails rather than recollection per se.

In a similar vein, conflicting conclusions regarding whether eye movements reflect conscious or unconscious influences of memory (Hannula, 2010; Ryan et al., 2000; Smith et al., 2006; Smith & Squire, 2008, 2017) may also be related to a failure to consider memory strength. For example, in most of these previous studies, conscious memory was assessed using dichotomous old/new recognition measures in which items falling above the participants' response criterion are treated as conscious (i.e., the hits), whereas items falling below the criterion are treated as unconscious (i.e., misses). However, because familiarity strength can vary, it is not clear if eye movements associated with misses reflect truly unconscious memory outside of awareness, or simply low levels of familiarity. In order to tease apart these possibilities, familiarity strength needs to be assessed using confidence-based measures to determine whether eye movement changes are related to conscious or unconscious memory.

Although confidence-based methods have been used in the memory literature in the past to examine unconscious memory, examining it alongside recollection and familiarity may be particularly important for understanding attention given the debates that have arisen regarding all three processes, outlined above. Specifically, it is possible that recollection, familiarity strength, and unconscious memory may each be characterized by unique relationships with viewing behavior, which would not be possible to assess without directly isolating them. For example, when only recollection and familiarity are examined, unconscious effects could be mistakenly attributed to familiarity, whereas when only unconscious and conscious memory are examined, familiarity effects could be mistakenly attributed to unconscious memory. However, to our knowledge, no study of attention during encoding and retrieval has examined familiarity and unconscious memory, or recollection and unconscious memory, in conjunction.

As mentioned above, another aspect of the relationship between eye movements and memory that has yet to be examined is the extent to which the spatial allocation of visual attention during encoding and retrieval, rather than general oculomotor measures (i.e., number of fixations and saccade amplitude), is related to different memory processes. This distinction is particularly important in light of findings that recollection and unconscious memory were each related to unique indices of spatial attention during search—but general oculomotor measures were not able to dissociate these memory processes (Ramey, Henderson, & Yonelinas, 2019). Importantly, the studies outlined above that examined saccade amplitude interpreted the relationship between decreased saccade amplitude during encoding and subsequent recollection to indicate that recollection is related to making less dispersed fixations (Kafkas & Montaldi, 2011; Sharot, Davidson, Carson, & Phelps, 2008)—that is, constraining viewing to a small scene region. However, because the spatial locations of fixations were not considered, it is not clear how recollection and familiarity may relate to the dispersion of attention across a stimulus during encoding. Furthermore, it is not yet known how spatial dispersion of attention during encoding and retrieval may relate to memory performance for scenes in general, in addition to the question of underlying memory processes.

Current Research

In the present study, we addressed these questions by examining how spatial eye movement patterns during encoding and retrieval of scenes are related to recollection, familiarity, and unconscious memory for those scenes. To this end, we used a recently developed confidence-based memory assessment method that allowed us to isolate the effects of these different memory processes (Ramey et al., 2019). We also developed two new eye movement measures of spatial attention to quantify resampling and dispersion of attention across a stimulus.

In two experiments, participants viewed a series of real-world scenes while their eye movements were tracked. The second experiment served as a replication of the first, using a different stimulus set and a different group of participants. During an initial study phase, participants viewed a series of scenes in two encoding tasks. In one encoding task, participants were asked to memorize each scene, whereas in the other, they were asked to judge each scene for its aesthetic appeal. The two different encoding blocks were included to test the generalizability of any effects obtained as well as to verify that the effects were not limited to conditions in which participants intentionally encoded the scenes. During a subsequent test phase, participants viewed the same scenes that they had viewed during the study phase (i.e., old scenes) along with randomly intermixed new scenes, and were asked to provide a recognition judgment for each scene. Memory awareness was measured by asking participants to rate memory confidence for each scene on a 6-point scale during the recognition judgment. Participants were told that if they could consciously recollect some qualitative aspect of the initial learning event, such as what they thought about when the scene was encountered earlier, they should respond “Recollect old (6);” otherwise, they rated their memory confidence by responding “I’m sure it’s old (5),” “Maybe it’s old (4),” “I don’t know (3),” “Maybe it’s new (2),” or “I’m sure it’s new (1).”

To isolate eye movement patterns that were related to recollection, we examined the old scenes that were confidently recognized as old and for which participants reported being able to retrieve specific details about the study event (i.e., “recollect old”). To assess eye movement patterns related to familiarity, we examined intermediate levels of memory confidence for old scenes, ranging from high familiarity strength (i.e., “I’m sure it’s old”) to low familiarity strength (i.e., “I’m sure it’s new”). To isolate eye movement patterns related to unambiguously unconscious memory, we examined the old scenes that participants were confident had not been studied (i.e., receiving a response of “I’m sure it’s new”). That is, examining only “sure new” responses ensured that we excluded any scenes for which there was even a weak sense of conscious memory (i.e., the “maybe it’s old,” “I don’t know” and “maybe it’s new” scenes).

Thus, any observed eye movement differences between new and old scenes given a response of "I'm sure it's new" could not be attributable to conscious memory.

In addition to employing a recent method of isolating recollection, familiarity, and unconscious memory, recent computational advances have allowed us to develop a new measure to directly assess the allocation of visual attention across a stimulus using the spatial distribution of eye movements. As outlined above, the measures that have been the primary focus of past studies of eye movements during encoding and retrieval are *number of fixations* and *saccade amplitude* (Kafkas & Montaldi, 2011, 2012; Sharot et al., 2008), which provide useful information about general oculomotor behavior but do not capture the deployment of visual attention across a stimulus. For example, an increase in the number of fixations does not necessarily entail more dispersed visual attention, as fixations are often clustered in a small region. Given that visual processing and successful encoding each heavily relies on where attention is deployed (e.g., Henderson & Hollingworth, 1999, 2003; Hollingworth, 2006), directly assessing the spatial distribution of viewing may be the key to understanding the relationship between visual attention and memory. Thus in addition to assessing previously used measures of oculomotor behavior such as saccade amplitude and number of fixations, we introduce two new measures of attention: one that uses cluster analysis to quantify spatial dispersion of eye movements across a stimulus, and one that provides a continuous measure of resampling of regions between viewings.

Given the conflicting findings surrounding the relationship between memory processes and eye movements during encoding and retrieval, a variety of outcomes are of interest. First, it is possible that the spatial allocation of visual attention via eye movements may selectively relate to recollection. This would be consistent with proposals that recollection is particularly related to visual attention, compared to familiarity or unconscious forms of memory (Holm & Mantyla, 2007; Kafkas & Montaldi, 2011, 2012; Mantyla & Holm, 2006). It would also indicate that attentional processes interact with some forms of memory but not others during encoding and retrieval. Alternatively, spatial attention may be associated with familiarity and/or unconscious forms of memory, suggesting that attentional processes interact more broadly with memory during encoding and retrieval. In addition, different memory processes may be related to different patterns of eye movements during encoding and/or retrieval—analogous to our prior findings using a visual search task (Ramey et al., 2019)—indicating that different types of memory are related to attention in different ways.

General Method

In two experiments, participants' eye movements were tracked while they viewed a series of real-world scene photographs during encoding (i.e., study phase) and retrieval (i.e., test phase). In the test phase, recognition memory was assessed for each scene, allowing us to examine how eye movements during encoding and retrieval related to recollection, familiarity, and unconscious memory for scenes. The second experiment served as a replication of the first, using different scenes and participants.

Experiment 1

Participants. Twenty-two undergraduates from the University of California, Davis completed the experiment for course credit. The sample size was selected to provide more than 80% power to detect the weakest effect of recollection on eye movements obtained in our prior study (Ramey et al., 2019). All participants had normal or corrected-to-normal vision, and provided informed consent in accordance with the study protocol as approved by the university IRB. The quality of each participant's eyetracking data was assessed by computing the mean percent signal across all trials, to determine whether there was excessive track loss due to blinks or calibration loss. All participants had greater than the preselected criterion of 75% signal (*M* = 96.5%) (Henderson & Hayes, 2017), such that they lost less than 25% signal; all participants were thus retained for analysis.

Stimuli. Stimuli were 204 photographs of real-world indoor scenes. All scenes were presented in color at 1024x768 pixels subtending a visual angle of approximately 25°x19° at presentation, and were free of people, animals, and text. Of these 204 scenes, 4 were used in practice trials, 150 were presented at study and test, and 50 were presented only at test. Stimulus presentation was counterbalanced, such that each scene appeared in different conditions (i.e., in one of the two study tasks, or as a new lure during test; see procedure) for different participants, to mitigate stimulus effects.[4]

Apparatus. Participants' eye movements were recorded using an SR Research EyeLink 1000+ tower mount eyetracker, sampling at 1000hz. A forehead and chin rest were used to reduce head movements, and eye movements were recorded from one eye though viewing was binocular. Stimuli were displayed on a monitor 85cm from the eyetracker, and the experiment was controlled with SR Research Experiment Builder software (SR Research, 2010a).

---

[4] Because there were 75 old scenes in each study task and 50 new scenes, the scenes were not perfectly counterbalanced between old and new conditions—we instead used a method similar to our prior study using the same memory contrasts (Ramey et al., 2019). This was a result of the time constraints imposed by eyetracking methods (e.g., dry eyes and blinking that worsens with time): We strategically included more old scenes for the analyses that would benefit from the additional power (i.e., familiarity strength and recollection). Effects related to unconscious memory required fewer new scenes to achieve the same power due to the relatively large numbers of "sure new" responses to new scenes, as in our prior study (e.g., Ramey et al., 2019).

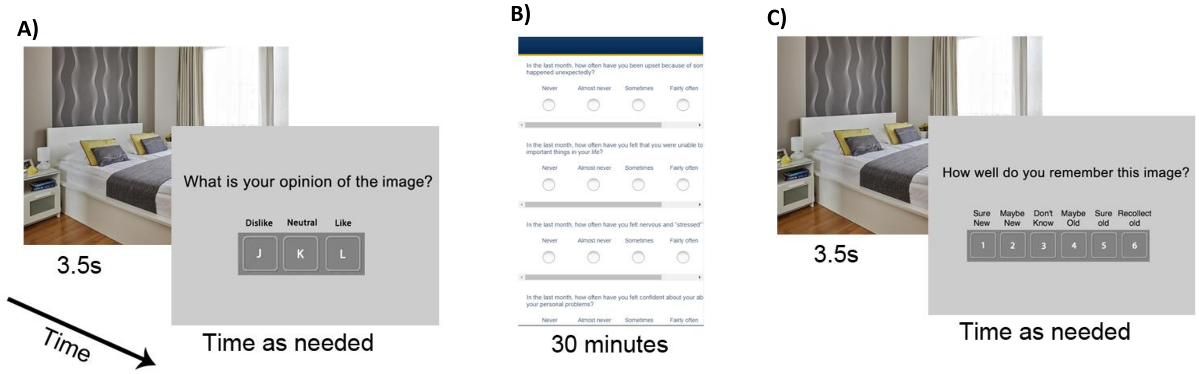


Figure 1. Illustration of the procedure. A) Study phase. Half of the scenes were presented in an aesthetic judgment task (i.e., participants were instructed to judge the image aesthetically and rate it as “dislike,” “neutral,” or “like”), whereas the other half were presented in a memorization task (i.e., participants were instructed to memorize the image and rate it as “not memorable,” “neutral”, or “memorable”). B) Delay between study and test, during which participants completed unrelated questionnaires. C) Test phase in which participants rated their recognition confidence.

Procedure. The experiment lasted 1.5 hours, and consisted of a study phase followed by a filled 30min delay, as well as a subsequent test phase (see Figure 1). Eye movements were recorded throughout the study and test phases. In both phases, each trial (i.e., each scene presentation) was preceded by a central fixation cross. Participants were given breaks every 50 trials and between phases, and the eyetracker was recalibrated after each of these breaks.

Study phase. During the study phase, participants were presented with 150 unique scenes split into two task blocks: an aesthetic judgment task and a memorization task. These tasks were selected to ensure that any effects obtained during encoding were not a product of a given task, but rather generalized across tasks (as prior work has shown that eye movements vary systematically between tasks; Castelhano, Mack, & Henderson, 2009; Henderson, Shinkareva, Wang, Luke, & Olejarczyk, 2013; Kardan, Berman, Yourganov, Schmidt, & Henderson, 2015; Mills, Hollingworth, Van der Stigchel, Hoffman, & Dodd, 2011). The order of the tasks was counterbalanced such that half of the participants completed the aesthetic judgment task first, whereas the other half completed the memorization task first. In each task, 75 scenes were presented for 3.5s each, allowing for an average of 12 fixations per trial. Each task was preceded by two practice trials to familiarize participants with the procedure.

In the aesthetic judgment task of the study phase, participants were asked to rate each scene based on how aesthetically pleasing they found it to be. Each trial consisted of a 3.5s scene presentation, followed by a grey response screen containing the prompt “What is your opinion of the photo?” as well as the key mappings for each response option. Responses were made on the keyboard, had no time limit, and consisted of “dislike,” “neutral,” and “like;” the response data were not used.

The memorization task of the study phase followed the same general procedure, but participants were instead asked to memorize the scenes. After each scene, they were asked to rate how memorable they found the scene to be. Participants were asked to give this response to ensure that the sequence of events in the memorization task was analogous to the aesthetic judgment task. Responses included “not memorable,” “neutral,” and “memorable;” again, the response data were not used.

Delay. Between the study and test phases, participants were moved to a computer in a different room to complete a 30min distractor task that included questionnaires (e.g., personality scales; see Appendix) that were not related to the present study.

Test phase. In the test phase, participants were presented with a series of scenes and asked to rate their memory for each scene. The test phase consisted of 200 trials: 150 old scenes, which had been presented in the study phase, and 50 new scenes, which had not been presented previously. Each scene was presented for 3.5s, as in the study phase, and was subsequently replaced by a recognition judgment screen.

For the recognition judgment, participants indicated whether or not they recognized the scene from the study phase. They were given as much time as they needed to select their response. Response options fell on a 1-5 and recollect scale made up of “sure new,” “maybe new,” “don’t know,” “maybe old,” “sure old,” and “recollect old” (Ramey et al., 2019; Yonelinas, 2002). Participants were instructed and tested on how to use this scale prior to beginning the test phase. A response of “recollect old” indicated that a participant could recall details of their experience of having seen the image in the study phase. Examples given to participants included remembering the study task in which they initially viewed the scene (i.e., memorization or aesthetic judgment), remembering an emotion they felt during prior exposure to the scene, and remembering ambient noise or sensations experienced during previous viewing of the scene. Participants were explicitly instructed that responses of “recollect old” and “sure old” were categorically different, rather than varying in memory strength. The other responses fell on a continuous gradient ranging from no memory to strong memory for an image, with a response of “sure old” indicating memory strength comparable to that of “recollect old” but without the additional episodic details.

Experiment 2

The methods and procedure were the same as those used in Experiment 1, with the following exceptions.

Participants. Forty-five undergraduates from the University of California, Davis completed the experiment for course credit. The sample size was selected to provide more than 98% power to detect the weakest effect of recollection on eye movements obtained in our prior study (Ramey et al., 2019), and more than 90% power to detect the weakest recollection effect obtained in Experiment 1. Eyetracking signal was greater than 75% in all participants (*M* = 94.7%), so all subjects were retained for analysis.

Stimuli. Stimulus characteristics were similar to Experiment 1, with the exception of content: rather than using only indoor scenes, a mix of new indoor and outdoor scenes were used. We also expected that the increased diversity of scenes would lead to higher recognition accuracy and more "recollect" responses, which would give us more power to detect recollection-specific effects.

Data Reduction and Analysis

Measures. Fixations and saccades were segmented with EyeLink's standard algorithm using velocity and acceleration thresholds (30°/s and 9500°/s^2; SR Research, 2010b). Eye movement data were imported offline into Matlab using the EDFConverter tool. We computed the eye movement measures as follows.

***Dispersion*.** The extent to which eye fixations are focused on a small number of regions or dispersed across a large number of regions of an image has been suggested to play an important role in how well an image is remembered, and which memory processes are involved (Kafkas & Montaldi, 2011, 2012; Sharot et al., 2008). To quantify attentional dispersion, we calculated the number of regions that were fixated in a scene, as well as the distance between those regions. This allowed us to determine how eye movements were spatially distributed, by taking into account how many fixation clusters were formed and how those clusters were distributed across the scene.

Specifically, we used a clustering algorithm to group fixations based on their spatial relation to each other, and created a composite dispersion score based on both the number of fixation clusters in a trial and the distance between those clusters. To this end, we first submitted the fixation locations for each trial, in terms of the (*x*,*y*) coordinates recorded by the eyetracker (with each coordinate reflecting one pixel in the 768x1024 pixel screen), to a k-means clustering algorithm (Maechler, 2015). Then, using a silhouette algorithm that assessed how similar each fixation was to its own cluster versus other clusters (Rousseeuw, 1987), we identified the optimal number of clusters that described the fixation locations (the minimum number of clusters was two). Using the optimal clustering solution, we then computed the average distance between the centroids of those clusters. Lastly, we created the dispersion score by multiplying the number of clusters by the average distance between the clusters, to capture the extent to which eye fixations were distributed across the display (Figure 2a-f). Therefore, higher values indicate that fixations were more distributed across a scene. This dispersion score was used in subsequent analyses, and we refer to it as dispersion.

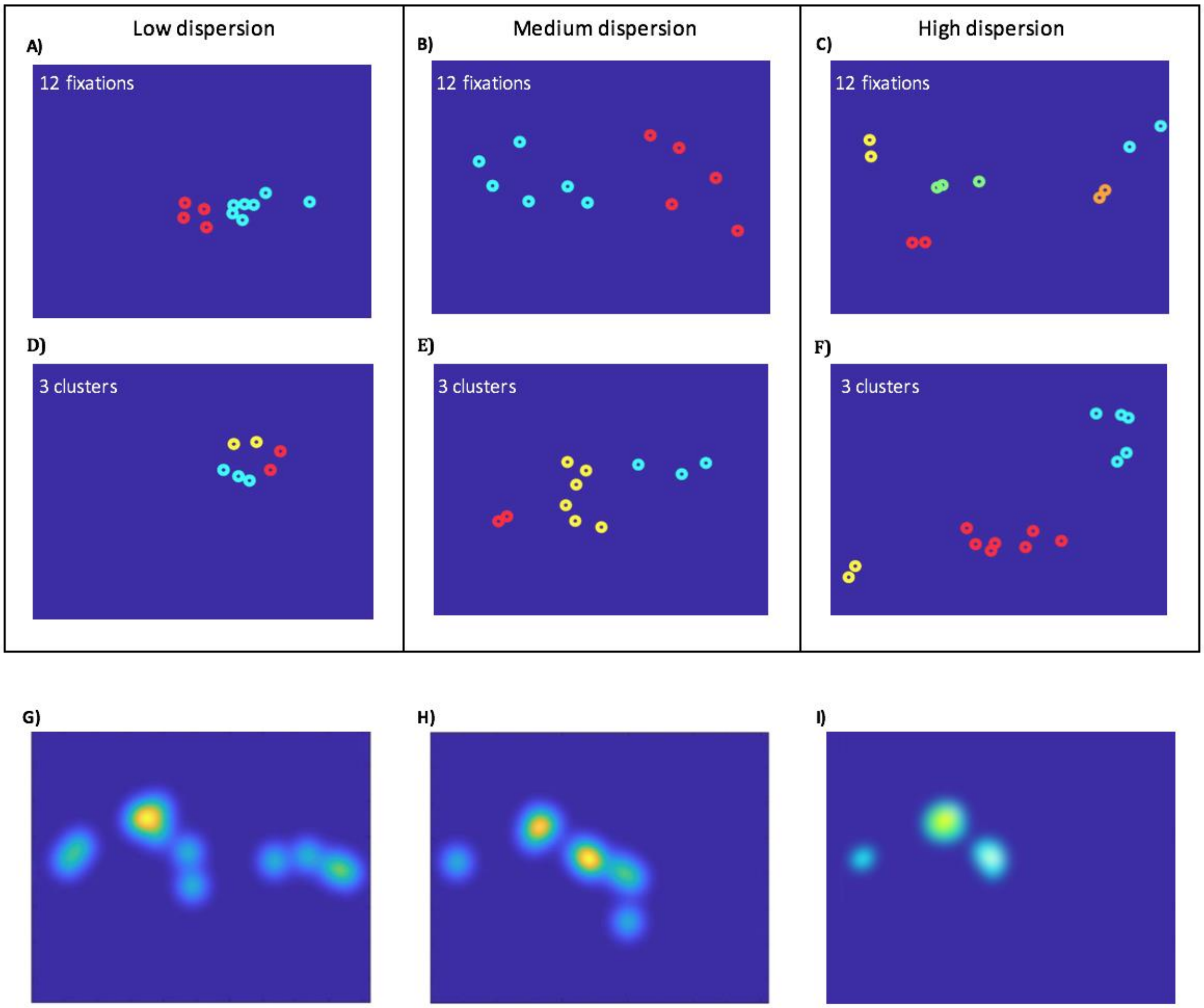


Figure 2. Visualizations of primary eye movement measures of interest. The rings in A-F represent fixations. Each cluster is denoted by a different color, such that the color of a fixation indicates its cluster membership. A) An example of a low-dispersion trial. B) An example of a medium-dispersion trial. C) An example of a high-dispersion trial. A-C are trials comprised of 12 fixations, which was the mode number of fixations, to illustrate how dispersion can vary for a given number of fixations. Similarly, D-F each contain three clusters of fixations, to illustrate how dispersion can vary for a given number of clusters. D) An example of a low-dispersion trial. E) An example of a medium-dispersion trial. F) An example of a high dispersion trial. G) An example of a smoothed fixation map, presented as a heatmap, of one subject viewing an image at study. H) The fixation map for that same subject and image during the test phase. I) The regions that overlap between the study and test maps, for the purposes of illustration. (The resampling measure captures the correlation between G and H.)

Resampling. To determine the extent to which the same regions were viewed between the study and test presentations of a scene, we created maps of the regions visited in each trial

(Figure 2g-h). To create these maps, we generated a matrix of fixation locations for each trial, in terms of (*x*,*y*) coordinates recorded by the eyetracker. A Gaussian low-pass filter with a cutoff frequency of -6dB (Henderson & Hayes, 2017) was then applied to the matrices to account for the fact that only the single pixel at the center of fixation is recorded by the eyetracker, whereas visual acuity is more diffuse (Bylinskii, Judd, Oliva, Torralba, & Durand, 2018). The resulting map represents the density of fixations at each pixel in a scene (Figure 2g-h). For a given participant, we computed the Pearson correlation coefficient between the fixation map generated while they viewed a scene at study with the fixation map of that same scene at test (Figure 2i). The resulting *resampling* value reflects the extent to which fixation locations were similar between study and test, such that higher values indicate that there was more overlap in the regions visited between study and test of a scene. It should be noted that resampling captures the proportion of fixations at retrieval that were devoted to revisiting previously viewed regions, such that it controls for the number of fixations made.

Additional measures. To better interface with prior literature, we also examined other, converging measures of eye movement behavior that have been used in past eyetracking studies. First, we assessed the number of regions visited in a trial, which has been taken to represent the dispersion of viewing. However, it should be noted that this measure does not consider the distance between regions, and a high number of regions visited could result from viewing constrained to a relatively small portion of the scene. For this analysis, we divided each scene into 64 evenly spaced rectangles (16x12 pixels), each of which defined a region. We then determined how many unique regions were fixated in a trial.

In addition to the number of regions visited, we also examined two eye movement measures that did not take spatial information into account. First, we calculated the number of fixations made in a trial. Second, we assessed saccade amplitude, which is the average distance between fixations in a trial. Each of these measures has also been considered by some to serve as a proxy for dispersion of viewing (Kafkas & Montaldi, 2011, 2012; Sharot et al., 2008).

Statistical models. Statistical analyses were conducted using linear mixed effects models with crossed random effects of participant and image, which allowed us to harness trial-by-trial (i.e., within-subjects) data while controlling for individual differences and stimulus effects. The models were estimated using the lmerTest package in R (Kuznetsova, Brockhoff, & Christensen, 2017), and were fit using maximum likelihood. The degrees of freedom and t values used were output by the linear mixed effects model for the variables of interest. The degrees of freedom were computed using the Satterthwaite approximation, and were rounded to the nearest integer in the manuscript. The models for each analysis were specified by regressing the eye movement measure in question on the memory variable, which depended on the type of memory being assessed: conscious recollection, unconscious memory, and familiarity strength (Table 1). Effect sizes were calculated as classical Cohen's *d*, as $2t/\sqrt{df}$ (Rosenthal & Rosnow, 1991), for the recollection and unconscious memory models, and as a standardized regression coefficient (β) for the linear gradient of familiarity strength.

Table 1

Linear Mixed Effects Model Specifications for Each Analysis

Fixed effect: Memory contrast	**Random effects**	**Included scenes**
Familiarity strength: “Sure new” through “Sure old”	Image, participant	All old scenes except those given a “recollect” response
Recollection: “Recollect” versus “Sure old”	Image, participant	All old scenes given responses of “recollect” or “sure old”
Unconscious memory: “Sure new” old scene versus “Sure new” new scene	Image, participant	All scenes, both old and new, given a “sure new” response

Note: In each model, the outcome was the eye movement measure of interest (e.g., dispersion, resampling). The fixed effect was the memory response given to a scene, or, in the case of unconscious memory, the old versus new status of the scene. Random effects were selected *a priori* to control for potential confounding influences of participant and scene, given the repeated measures design.

Results

Memory Accuracy

In Experiment 1, the percentage of scenes that received a recognition confidence response of “recollect,” “sure old,” “maybe old,” “don’t know,” “maybe new,” “sure new,” respectively, were 34%, 19%, 16%, 15%, 11%, and 5% for old scenes, and 3%, 5%, 15%, 21%, 27%, and 29% for new scenes. In Experiment 2, the percentage of scenes receiving these respective responses were 46%, 25%, 11%, 8%, 6%, and 4% for old scenes, and 2%, 3%, 8%, 13%, 27%, and 47% for new scenes. These results suggest that participants were able to discriminate between old and new scenes, and used the full range of response options. Furthermore, as expected, Experiment 2 yielded more recollection responses to old scenes, and higher recognition accuracy overall.

Task Effects

It is possible that the two tasks that participants performed while studying the scenes (i.e., memorization and aesthetic judgment) yielded fundamentally distinct relationships between memory responses and the eye movement measures examined. To address this possibility, each eye movement measure was regressed on the interaction between memory response and the task performed at study. There were no significant interactions between study task and memory response at study or test, for dispersion or resampling (*p*s >.12). Therefore, subsequent analyses collapsed across task.

Dispersion at Study Predicting Subsequent Memory

Familiarity strength. To assess how the spatial distribution of attention during the study phase predicted subsequent memory for the scenes, we compared dispersion of fixations between scenes that were given different memory responses in the test phase (for a review of the subsequent memory method, see Paller & Wagner, 2002). We first examined the effects of dispersion on familiarity strength by assessing whether there was a linear relationship between dispersion at study and subsequent familiarity confidence, using old scenes that were given any response except "recollect old" (Table 1). In both studies, we found that higher dispersion during study predicted higher subsequent familiarity strength, (Exp 1: $\beta = .10$, $t(2110) = 4.90$, $p < .0001$; Exp 2: $\beta = .07$, $t(3575) = 4.01$, $p < .0001$; Figure 3a-b). This suggests that more dispersed viewing during encoding of a scene, such that attention is more distributed across the scene, leads to subsequently higher familiarity.

Recollection. We then examined the extent to which dispersion of attention predicted subsequent recollection, by comparing scenes that were subsequently endorsed as "sure old" (i.e., high-confidence familiarity) with scenes subsequently endorsed as "recollect old" (Table 1). In both experiments, dispersion was numerically higher for recollected than for high-confidence familiar scenes, but the difference only reached statistical significance in Experiment 1, *t*(1734) = 2.37, *p* = .02, *d* = 0.11 (Experiment 2: *t*(4584) = 0.67, *p* = .50, *d* = 0.02). This suggests that more dispersed viewing may lead to a slight increase in recollection above high-confidence familiarity under some conditions, but the effect is not as consistent as the effect of dispersed viewing on familiarity strength.

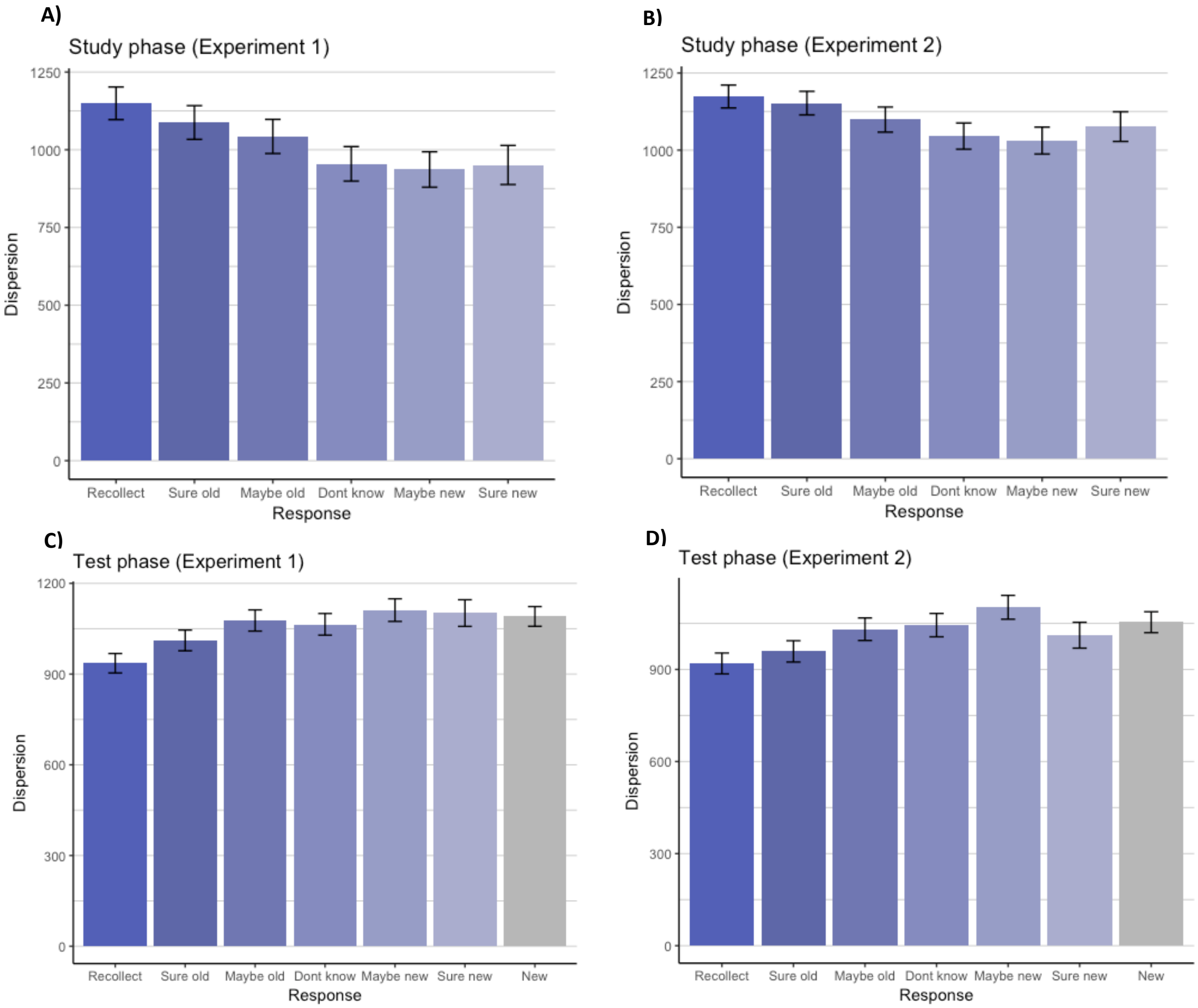


Figure 3. Dispersion at study and test by memory response. Estimated marginal means controlling for participant and image are plotted, and the error bars represent the standard error of these estimated means from the model. A) Dispersion during the study phase in Experiment 1, sorted by subsequent memory response. B) Dispersion during the study phase in Experiment 2. C) Dispersion during the test phase in Experiment 1. The “new” bar includes scenes that were new in the test phase, and therefore contained no memory. All other bars besides “new” only include old scenes. D) Dispersion during the test phase in Experiment 2.

Dispersion at Test Related to Memory

Familiarity strength. We examined dispersion of viewing during the test phase, in which participants inspected each scene to determine if they recognized it. To assess how dispersion during the test phase related to familiarity strength, we compared dispersion between scenes that had been viewed in the study phase across different memory responses ranging from “sure new” through “sure old,” as in the model used for the study phase data (Table 1). Dispersion decreased significantly as familiarity strength increased in both Experiment 1, $\beta = -.06$, $t(2109) = -2.94$, $p = .003$,and Experiment 2, $\beta = -.08$, $t(3534) = -4.92$, $p < .0001$. These results suggest that less dispersed viewing during retrieval is related to increased familiarity strength.

To ensure that these effects were unique to old scenes (i.e., scenes for which participants had memory), we submitted the new scenes to the same analysis. Dispersion was not significantly related to subjective experiences of familiarity in these new scenes in Experiment 1, $\beta = -.05$, $t(1054)= -1.73$, $p = .08$, nor Experiment 2, $\beta = .02$, $t(2191)= 0.97$, $p > .25$. This suggests that dispersion was not simply related to memory responses in the absence of true memory. However, we note that the familiarity strength effect in old scenes was driven largely by the higher confidence familiarity responses (i.e., “sure old” and “maybe old” responses), and there were many fewer new scenes associated with this level of confidence.

Recollection. To assess how recollection was related to dispersion of attention in the test phase, we again compared dispersion between scenes given a response of “recollect” and scenes given a response of “sure old” (Table 1). Similar to the effects of familiarity strength, recollection was related to a reduction in dispersion in both Experiment 1, $t(1735) = -4.01$, $p < .0001$, $d = -0.19$, and Experiment 2, $t(4589) = -4.36$, $p < .0001$, $d = -0.13$. It therefore appears that both recollection and familiarity are related to less dispersion during retrieval. Similar to familiarity strength, there was no significant relation between dispersion and recollection in new scenes in Experiment 1, $t(60) = -0.15$, $p > .25$, nor in Experiment 2, $t(106) = -0.17$, $p > .25$.

Unconscious memory. To assess unconscious memory, we compared scenes that were previously viewed with scenes that were newly presented in the test phase.[5] Importantly, we only examined scenes that were given a response of "sure new," indicating that participants were confident that they had not seen them before (Table 1). This strict criterion for unconscious memory (i.e., only considering scenes that were endorsed as "sure new" rather than all misses) ensured that none of the scenes used in the unconscious memory contrast were contaminated by conscious recollection or familiarity, and that the scenes differed only in terms of whether or not the participant had seen them previously. Dispersion did not relate to unconscious memory in either Experiment 1, $t(397) = -0.44$, $p = .66$, $d = -0.04$, or Experiment 2, $t(1236) = 0.04$, $p = .97$, $d = 0.00$. Furthermore, Bayesian analysis provided substantial evidence for this null effect in Experiment 1, $BF_{10} = 0.13$, and strong evidence for the null in Experiment 2, $BF_{10} = 0.07$, such that any numerical difference was more than six times more likely to be explained by chance than by unconscious memory.[6] This suggests that unconscious memory did not influence dispersion of viewing.

Resampling Scene Regions Between Study and Test

---

[5] It should be noted that unconscious memory could only be examined for eye movements made during the test phase, and not in the study phase, because it is not possible to have a comparison group of non-studied scenes in the study phase. Similarly, it was not possible to assess how resampling related to unconscious memory, because there is no direct way to obtain a resampling value for a new scene (i.e., a scene that was only viewed once).

[6] By convention, a $BF_{10} < 0.33$ indicates substantial evidence for the null hypothesis, and a $BF_{10} < 0.10$ indicates strong evidence for the null hypothesis (Jeffreys, 1961).

Familiarity strength. To assess how visiting scene regions at test that were initially visited at study related to memory for a scene, we compared the degree of resampling between scenes given different memory responses. Higher resampling scores reflect both increased resampling of previously visited regions, and as a corollary, decreased sampling of new regions. We first examined familiarity strength, by assessing whether resampling was linearly related to familiarity strength (Table 1). Resampling was significantly higher for scenes that were more familiar in both Experiment 1, $\beta = .10$, $t(2064) = 4.59$, $p < .0001$, and Experiment 2, $\beta = .14$, $t(3543) = 8.12$, $p < .0001$, which suggests that revisiting the same regions between study and test is associated with increased familiarity for a scene.

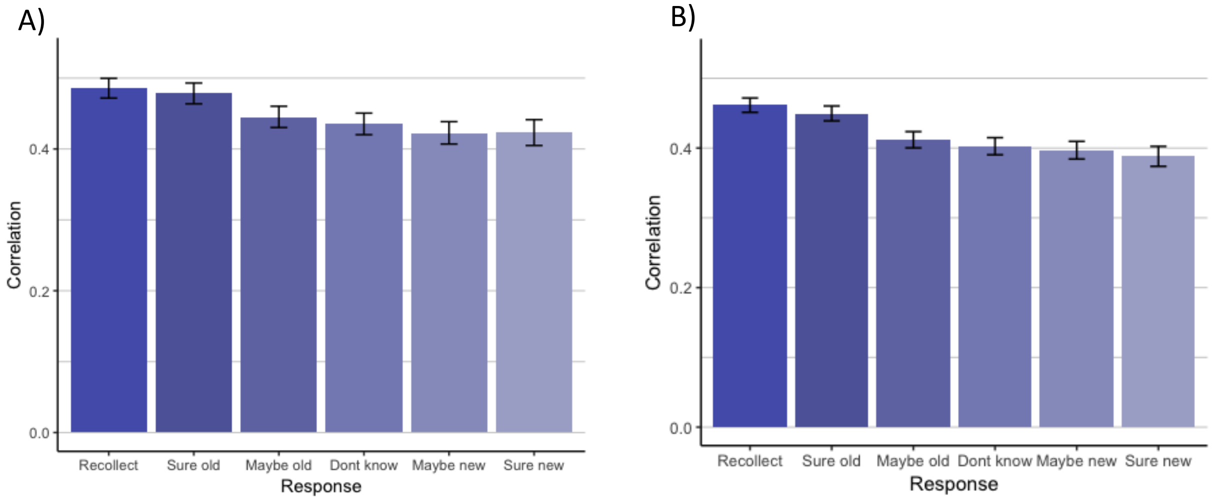


Figure 4. Degree of resampling of scene regions between study and test by memory response. Resampling is given as the correlation between the fixation density maps at study and test. Estimated marginal means controlling for participant and image are plotted, and the error bars represent the standard error of these estimated means from the model. A) Resampling in Experiment 1. B) Resampling in Experiment 2.

Recollection. Resampling was then compared between scenes endorsed as “recollect old” and scenes endorsed as “sure old” (Table 1), to determine if resampling was related to recollection. Resampling was numerically higher for recollection than for high-confidence familiarity in both experiments, but the effect only reached statistical significance in Experiment 2, *t*(4098) = 3.05, *p* = .002, *d* = 0.10 (Experiment 1: *t*(1643) = 0.38, *p* = .7, *d* = 0.02).

Additional Analyses

In order to facilitate comparisons to previous studies, we examined several commonly used eye movement measures with respect to the current recollection, familiarity, and unconscious memory contrasts, and present the results of these analyses in Table 2. In general, these measures led to results that were similar to those obtained using the dispersion measure, which is also provided in Table 2 for ease of comparison between measures. Additionally, we conducted a model comparison to determine if the presently developed eye movement measures (i.e., dispersion and resampling) predicted memory above and beyond previously used measures, and found that a model with dispersion and resampling was superior in both experiments (see Appendix). Lastly, we re-ran the analyses examining dispersion and unconscious memory by collapsing all misses, to ensure that the lack of effects was not due to our strict definition of misses (see Appendix).

Table 2

Memory Effects for Eye Movement Measures Related to Dispersion of Attention

Phase	Measure	Exp.	Fam. β	Fam. *p*	Rec. *d*	Rec. *p*	Uncon. *d*	Uncon. *p*
Study								
	Sac. Amplitude	1	**.07**	**<.001**	-0.04	.43	-	-
	Sac. Amplitude	2	**.05**	**<.001**	-0.05	.12	-	-
	# fixations	1	**.11**	**<.0001**	**0.15**	**.002**	-	-
	# fixations	2	**.11**	**<.0001**	**0.13**	**<.0001**	-	-
	# regions visited	1	**.10**	**<.0001**	0.05	.31	-	-
	# regions visited	2	**.10**	**<.0001**	**0.09**	**.002**	-	-
	Dispersion	1	**.10**	**<.0001**	**0.11**	**.02**	-	-
	Dispersion	2	**.07**	**<.0001**	0.02	.50	-	-
Test								
	Sac. Amplitude	1	.01	.6	-0.06	.21	-0.07	.46
	Sac. Amplitude	2	**-.06**	**<.001**	**-0.16**	**<.0001**	0.02	.71
	# fixations	1	**.04**	**.03**	**-0.24**	**<.0001**	-0.02	.88
	# fixations	2	-.02	.16	**-0.13**	**<.0001**	0.04	.47
	# regions visited	1	-.02	.44	**-0.22**	**<.0001**	-0.07	.49
	# regions visited	2	**-.09**	**<.0001**	**-0.19**	**<.0001**	0.05	.35
	Dispersion	1	**-.06**	**.003**	**-0.19**	**<.0001**	-0.04	.66
	Dispersion	2	**-.08**	**<.0001**	**-0.13**	**<.0001**	0.00	.97

Note: Exp. = Experiment; Fam. = Familiarity; Rec. = Recollection; Uncon = Unconscious; Sac. Amplitude = Saccade Amplitude. See Table 1 for descriptions of models for each memory process. Significant results are bolded.

Saccade amplitude. To better assess how the present results fit with past findings (Kafkas & Montaldi, 2011, 2012; Sharot et al., 2008), we examined saccade amplitude, which is the distance between fixations. The relation between saccade amplitude and memory was less consistent across the present experiments than dispersion, such that some effects only emerged in Experiment 2. However, its relation with familiarity strength was more consistent than with recollection, and was in the same direction as dispersion. In contrast to the present results, a prior study reported that average saccade amplitude during encoding increased with familiarity strength, but was lower for recollection than high-confidence familiarity (Kafkas & Montaldi, 2011). One possible reason for the discrepant results is the fact that the prior study did not control for participant-level effects, whereas the present study did. In support of this hypothesis, when we removed the random effect of participant that we included to eliminate potentially confounding individual differences, we found that we were able to replicate the pattern of saccade amplitudes obtained by Kafkas and Montaldi (2011) during encoding: increased saccade amplitude predicted subsequent familiarity strength (ps <.0001 in both experiments), but *decreased* saccade amplitude predicted subsequent recollection (ps <.05 in both experiments). When participant-level effects were controlled for, on the other hand, only the familiarity strength effect remained (Table 2). In contrast, the presently developed measure of dispersion did not show a reversal of the recollection effect when the participant covariate was removed. This suggests that saccade amplitude effects that were previously attributed to trial-by-trial changes in eye movements may instead reflect a relation between individual differences in saccade amplitude (Henderson & Luke, 2014) and individual differences in recollection—whereas dispersion may be less subject to individual differences.

Number of fixations. The pattern of results with respect to the number of fixations in a trial was largely the same as that observed using the dispersion measure. Specifically, more fixations made at study predicted subsequent recollection and high familiarity strength, whereas fewer fixations at test were related to recollection and high familiarity strength (see Table 2).

Number of regions visited. To provide a converging measure with the presently developed measure of spatial dispersion, we computed the number of regions visited. The results largely paralleled those obtained using the dispersion measure, such that familiarity strength was related to the number of regions visited, with the exception of Experiment 1 at retrieval. Interestingly, while dispersion at encoding was related to recollection in Experiment 1 but not Experiment 2, the opposite was true for number of regions visited.

See the Appendix for supplemental measures, correlations between eye movement measures, and analyses ensuring that the pattern of results was not altered by the differences in accuracy rates and scene type between experiments, or the criterion for unconscious memory.

Discussion

In the present study, we examined how recollection, familiarity strength, and unconscious memory related to the deployment of spatial visual attention during encoding and retrieval of real-world scenes. Participants' eyes were tracked as they viewed a series of scenes during a study and test phase, and they provided confidence-based recognition judgments for each scene during the test phase. Recollection, familiarity strength, and unconscious memory were isolated based on the recognition judgments, and three different sets of analyses were used to assess how eye movement patterns related to these memory processes. In two experiments, we found that familiarity strength was robustly associated with viewing patterns during both encoding and retrieval. Specifically, we found that a more dispersed distribution of eye movements during encoding predicted subsequently stronger familiarity, whereas less dispersed eye movements during retrieval were related to stronger familiarity. These effects also emerged in additional measures of general eye movement behavior (i.e., saccade amplitude, number of fixations, and number of regions visited), indicating that there is a consistent relationship between familiarity strength and eye movement behavior. Furthermore, we found that increased resampling of previously viewed regions during retrieval was related to familiarity strength as well. Interestingly, we found that recollection followed the same trends as familiarity strength, with no evidence for effects unique to recollection-based memory. That is, like familiarity strength, recollection was also predicted by more dispersed eye movements at encoding, less dispersed eye movements at retrieval, and resampling of regions between encoding and retrieval; however, these effects were somewhat less reliable across studies. Moreover, we found no evidence for an influence of unconscious memory, such that no differences in eye movement patterns were observed between new scenes and old scenes for which participants did not have conscious memory (i.e., high-confidence misses).

Prior investigations of recollection and familiarity have concluded that the two processes fundamentally differ in their relationship with patterns of sampling behavior during both encoding and retrieval (Kafkas & Montaldi, 2011, 2012). Specifically, it has been suggested that recollection involves decreased saccade amplitude compared to familiarity during both encoding and retrieval (Kafkas & Montaldi, 2011; Sharot et al., 2008, but see Kafkas & Montaldi, 2012), and that overlap in regions visited between encoding and retrieval is uniquely related to recollection (Holm & Mantyla, 2007; Mantyla & Holm, 2006). These effects have been interpreted as showing that recollection and familiarity have qualitatively different relationships with eye movement behavior, such that recollection has a particularly strong relationship with eye movements to allow for the retrieval of specific details. However, it is possible that memory strength differences between recollection and familiarity may have confounded prior results, such that comparing recollection to all levels of familiarity strength—rather than strength-matched familiarity—may have obscured effects arising from differences in memory strength (Montaldi et al., 2006; Yonelinas et al., 2005). Taking memory strength into account, the present results reveal that familiarity strength is very strongly related to attentional dispersion, as well as other measures of eye movement behavior including saccade amplitude, across both encoding and retrieval. Moreover, we found that recollection and familiarity did not involve qualitative differences in dispersion of viewing or resampling of regions: The relationship of these eye movements with recollection paralleled their relationship with familiarity strength. Together, these results suggest that eye movement patterns during naturalistic encoding and retrieval may not clearly dissociate along the lines of recollection and familiarity.

The current results indicate that both familiarity and recollection were related to the dispersion of viewing and resampling of previously studied scene regions. Future work, however, will be needed to clarify the precise roles that these eye movement patterns play in memory, particularly during retrieval. For example, the eye movements observed during the retrieval phase could reflect either processes facilitating memory retrieval (e.g., resampling studied regions in a scene may increase study-test similarity and so may facilitate retrieval; Wynn et al. 2016), or they could reflect decision processes that occur after the memory is retrieved (e.g., increased resampling of studied regions could reflect evaluation of the strength or the quality of the specific details that were retrieved from memory; Holm and Mantyla 2007). Prior work has indicated that limiting eye movements during retrieval can significantly reduce recollection-based memory responses (Mantyla & Holm, 2006; Schwedes, Scherer, & Wentura, 2019; Schwedes & Wentura, 2019), suggesting that the presently observed eye movements may have been involved in facilitating recollection; however, we cannot rule out the possibility that they may have reflected additional post-retrieval processes as well. Similarly, the eye movement effects related to familiarity could also reflect pre-retrieval and/or post-retrieval processes. For example, there is evidence that familiarity is related to perceptual fluency, such that ease of identification of stimuli contributes to increased familiarity for those stimuli (Whittlesea, 1993). One possible explanation for the observed relationship between eye movements and familiarity in the present study, therefore, is that attention focused on previously visited regions during retrieval reflects more fluent visual processing and leads to the subjective experience of familiarity. But another possibility is that the eye movement effects primarily reflect post-retrieval processing driven by familiarity. Teasing apart these possibilities may be particularly useful in furthering our understanding of the interplay of attention and memory.

Given that eye movement changes resulting from experience are frequently considered to be underpinned by memory that occurs outside of awareness (e.g., Hannula, 2010; Hannula & Ranganath, 2009; Ryan, Althoff, Whitlow, & Cohen, 2000), it is perhaps surprising that we found no evidence that eye movements were related to unconscious memory in the present studies. One possible explanation for this discrepancy is that we simply lacked the statistical power to detect unconscious memory effects. However, this explanation seems unlikely for a number of reasons. First, even though power is a concern with a nonsignificant standard hypothesis test, the Bayes factors that we achieved in both experiments met the convention for confidence that the null results were not due to lack of power, but instead were more likely to have occurred under the null hypothesis. Second, we doubled the sample size in Experiment 2 such that we had more than 98% power to detect previously obtained effects of memory on eye movements—but again found no evidence for a difference between new and missed old scenes. Third, we ran an additional analysis collapsing “sure new” and “maybe new” trials into a single “all misses” category, thereby doubling the number of trials included in the analysis, and this further verified the results of the initial analysis (see Appendix). Fourth, inspection of the numerical differences between missed and new scenes in both Experiments 1 and 2 indicated that the results were reversed from what would be expected of an unconscious memory effect, given prior findings. Another potential explanation for the lack of effect is that unconscious memory may influence some eye movement measures, but not the specific eye movement measures we examined in the current study. However, we did examine a variety of measures (i.e., dispersion, number of fixations, saccade amplitude, number of regions visited), including those that have previously been thought to relate to unconscious memory (i.e., number of fixations; Althoff & Cohen, 1999; Ryan et al., 2000), but none of them showed any relation with unconscious memory. The present findings, therefore, may be useful in informing the debate surrounding whether changes in oculomotor behavior and visual attention at retrieval are related to conscious or unconscious memory: When unconscious memory was isolated from confounding influences of weak conscious memory, we found only conscious effects and substantial evidence against unconscious effects.

Despite the evidence for a lack of unconscious effects in this paradigm, it is possible that there are other tasks that are more appropriate for detecting unconscious memory effects. As one example, in a prior study we found that during memory-guided search (i.e., contextual cueing, which is an implicit memory task), the overall efficiency of participants’ scanpaths was influenced by unconscious memory, but not recollection or familiarity (Ramey et al., 2019). Thus, one possibility warranting further investigation is that unconscious memory effects on attention may be suppressed or masked under certain explicit retrieval conditions, such as those used in the present study. Although this is a relatively unexplored area, there is evidence that conscious and unconscious memory can compete for expression in some circumstances (Henson, Shallice, Gorno-Tempini, & Dolan, 2002). For example, in an implicit test of memory for faces (i.e., fame judgments), Henson and colleagues (2002) found that activity in the fusiform gyrus was reduced for faces that had been studied earlier, which is considered a marker of neural priming. In an explicit version of the task (i.e., recognition memory test), however, the neural priming effects were no longer observed but were replaced by medial temporal lobe activity—a marker of conscious, explicit memory. Thus, the current results should not be interpreted as ruling out the possibility that unconscious memory may impact some types of eye movements under some conditions, but they do suggest that under explicit retrieval conditions, eye movements and the allocation of visual attention are more tightly coupled with conscious memory.

The finding of a robust relation between familiarity strength and eye movement behavior has important implications for understanding the relationship between memory and attention more broadly. For example, building on prior results showing that making more fixations during encoding predicts better subsequent memory (Kafkas & Montaldi, 2011; Loftus, 1972), the present results suggest that dispersing those fixations broadly across a stimulus also predicts improved memory—regardless of the task performed during encoding (i.e., memorization or aesthetic judgment). Perhaps surprisingly, our findings at retrieval suggest that instead of widely distributing attention in a similar fashion to encoding, the opposite pattern is optimal: restricting fixations to a more constrained area of the stimulus at retrieval is related to better memory. Moreover, those constrained fixation patterns may function to focus attention on regions that were previously visited—as suggested by the relation between increased resampling of previously viewed regions (and therefore decreased sampling of new regions) and overall memory strength. Together, these results are broadly consistent with an account of attentional deployment wherein attention is widely distributed at encoding to facilitate maximal sampling of information, and focused in on the most relevant regions at retrieval to facilitate comparisons with internal memory representations.

The presently developed measures of dispersion and resampling may prove to be particularly useful new tools for understanding the relationship between memory and attention. First, the present resampling measure provides a continuous metric of resampling behavior that does not require arbitrary cutoffs, whereas past assessments of memory processes and resampling behavior have all utilized categorical measures (i.e., the proportion of the first 3 test fixations that were within 2 degrees of study fixations; Holm & Mantyla 2006; Holm & Mantyla 2007). Second, dispersion demonstrated a more consistent relationship with familiarity than did previously used measures such as number of fixations and saccade amplitude: Whereas dispersion was strongly related to familiarity strength in both experiments at both encoding and retrieval, the number of fixations and saccade amplitude were not reliable across experiments (Table 2), and dispersion outperformed prior measures in a model comparison (Appendix). Furthermore, dispersion was not influenced by subject and image effects, whereas the relationship between saccade amplitude and recollection was reversed depending on if subject effects were controlled for—suggesting that dispersion is more robust across statistical methods. The current measure of dispersion also has clear potential for applications beyond memory research, to questions of visual attention more broadly. For example, examining cluster-based dispersion may be useful for understanding the processes involved in perceptual discriminations such as change detection, given that change-related differences have been found in converging eye movement measures such as saccade amplitude (Henderson & Hollingworth, 2003). Furthermore, prior work indicates that saccade amplitudes and the number of fixations vary between emotional and neutral scenes (Bradley et al., 2011), and dispersion may provide a more robust, sensitive method of examining differences in attentional distribution for emotional stimuli. Therefore, given its stability and potential for applications beyond memory research, dispersion may prove to be a sensitive new index for visual attentional deployment that could be useful for attention researchers in general.

The present results may also have important implications for longstanding theoretical debates in both attention and scene memory, particularly with respect to the importance of stimulus properties versus cognitive factors in the control of attention and memory for scenes (e.g., Henderson, 2007; Van der Stigchel et al., 2009). For example, much of the research into how attention is controlled has been focused on predicting attention by quantifying various scene properties, such as the salience (Itti & Koch, 2000; Itti, Koch, & Niebur, 1998) and semantics (Henderson & Hayes, 2017, 2018) of different scene regions. Some dominant theories have even assumed that attention is controlled primarily by bottom-up visual features, with cognitive factors serving only to modulate the prioritization of visually salient regions (Henderson, 2007; Itti & Koch, 2001; Tatler, Hayhoe, Land, & Ballard, 2011). In contrast, a growing body of research has indicated that cognitive factors unrelated to scene properties, such as task goals, can guide attention independently of salience (Henderson, 2003, 2007; Tatler et al., 2011). The role of episodic memory as a source of cognitive guidance, however, has not yet been well-defined—and the present results suggest that subjective familiarity strength may comprise a unique form of attentional guidance warranting further investigation. In a similar vein, much research has been devoted to examining memory for scenes in terms of scene memorability: the intrinsic aspects of scenes that tend to elicit better memory (e.g., Bainbridge, Hall, & Baker, 2019). For example, some findings have indicated that scenes containing certain features such as faces tend to be more memorable (e.g., Isola, Parikh, Torralba, & Oliva, 2011), and that scenes eliciting greater consistency between different participants' viewing patterns tend to be more memorable (Mancas & Le Meur, 2013). In fact, the effects obtained in prior studies of memory and visual attention (e.g., Kafkas & Montaldi, 2011; Loftus, 1972; Sharot et al., 2008) could perhaps reflect differences in image properties: For example, certain scenes may be more memorable and also elicit more fixations, both due to inherent scene properties (e.g., having many interesting regions), thus leading to the observed relation between memory and an increased number of fixations. Given that we controlled for image effects, however, the present findings may be uniquely poised to contribute to our understanding of how scenes are remembered independently from scene properties. Specifically, the present findings are the first to our knowledge to point to a robust role of visual attention in scene memory strength that is disentangled from the influence of image properties.

Both the results and newly developed methods of the present study may prove useful in motivating future investigations in both attention and memory research. For example, the present findings combined with prior evidence for eye movement measures that separately index recollection and unconscious memory during search (Ramey et al., 2019) suggest that recollection, familiarity, and unconscious memory can each be indexed using eyetracking, which may be particularly useful for assessing memory processes in nonverbal and patient populations. Additionally, given the sensitivity of the presently developed measure of attentional dispersion to familiarity strength—and the fact that many cognitive processes may indirectly support memory (e.g., Johnson, 1992)—future research aimed at determining whether dispersion of attention might also index phenomena such as cognitive load, mind-wandering, or aspects of executive functioning may prove fruitful. Finally, if causal investigations show that manipulating attentional dispersion can increase familiarity strength, a dispersion-based intervention could perhaps be fruitfully applied to improve learning in a real-world setting.

References

Althoff, R. R., & Cohen, N. J. (1999). Eye-Movement-Based Memory Effect: A Reprocessing Effect in Face Perception. *Journal of Experimental Psychology: Learning Memory and Cognition*, *25*(4), 997–1010. https://doi.org/10.1037/0278-7393.25.4.997

Bainbridge, W. A., Hall, E. H., & Baker, C. I. (2019). Drawings of real-world scenes during free recall reveal detailed object and spatial information in memory. *Nature Communications*, *10*(1), 5. https://doi.org/10.1038/s41467-018-07830-6

Bylinskii, Z., Judd, T., Oliva, A., Torralba, A., & Durand, F. (2018). What do different evaluation metrics tell us about saliency models? *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1–24. https://doi.org/10.1109/TPAMI.2018.2815601

Bradley, M. M., Houbova, P., Miccoli, L., Costa, V. D., & Lang, P. J. (2011). Scan patterns when viewing natural scenes: Emotion, complexity, and repetition. *Psychophysiology*, *48*(11), 1544–1553. https://doi.org/10.1111/j.1469-8986.2011.01223.x

Castelhano, M. S., Mack, M. L., & Henderson, J. M. (2009). Viewing task influences eye movement control during active scene perception. *Journal of Vision*, *9*(3), 6–6. https://doi.org/10.1167/9.3.6

Foulsham, T., & Kingstone, A. (2013). Fixation-dependent memory for natural scenes: An experimental test of scanpath theory. *Journal of Experimental Psychology: General*, *142*(1), 41–56. https://doi.org/10.1037/a0028227

Hannula, D. E. (2010). Worth a glance: using eye movements to investigate the cognitive neuroscience of memory. *Frontiers in Human Neuroscience*, *4*, 166. https://doi.org/10.3389/fnhum.2010.00166

Hannula, D. E., & Ranganath, C. (2009). The eyes have it: hippocampal activity predicts expression of memory in eye movements. *Neuron*, *63*(5), 592–9. https://doi.org/10.1016/j.neuron.2009.08.025

Henderson, J. M. (2003). Human gaze control during real-world scene perception. *Trends in Cognitive Sciences*, *7*(11), 498–504. https://doi.org/10.1016/j.tics.2003.09.006

Henderson, J. M. (2007). Regarding Scenes. *Current Directions in Psychological Science*, *16*(4), 219–222. Retrieved from http://journals.sagepub.com/doi/pdf/10.1111/j.1467-8721.2007.00507.x

Henderson, J. M., & Hayes, T. R. (2017). Meaning-based guidance of attention in scenes as revealed by meaning maps. *Nature Human Behaviour*, *1*(10), 743–747. https://doi.org/10.1038/s41562-017-0208-0

Henderson, J. M., & Hayes, T. R. (2018). Meaning guides attention in real-world scene images: Evidence from eye movements and meaning maps. *Journal of Vision*, *18*(6), 10. https://doi.org/10.1167/18.6.10

Henderson, J. M., & Hollingworth, A. (1999). The role of fixation position in detecting scene changes. *Psychological Science*, 438–443. Retrieved from https://journals.sagepub.com/doi/pdf/10.1111/1467-9280.00183

Henderson, J. M., & Hollingworth, A. (2003). Eye movements and visual memory: Detecting changes to saccade targets in scenes. *Perception and Psychophysics*, *65*(1), 58–71. https://doi.org/10.3758/BF03194783

Henderson, J. M., & Luke, S. G. (2014). Stable individual differences in saccadic eye movements during reading, pseudoreading, scene viewing, and scene search. *Journal of Experimental Psychology: Human Perception and Performance*, *40*(4), 1390–1400. https://doi.org/10.1037/a0036330

Henderson, J. M., Shinkareva, S. V., Wang, J., Luke, S. G., & Olejarczyk, J. (2013). Predicting Cognitive State from Eye Movements. *PloS ONE*, *8*(5). https://doi.org/10.1371/journal.pone.0064937

Henson, R. N. A., Shallice, T., Gorno-Tempini, M. L., & Dolan, R. J. (2002). Face repetition effects in implicit and explicit memory tests as measured by fMRI. *Cerebral Cortex*, *12*(2), 178–186. https://doi.org/10.1093/cercor/12.2.178

Hollingworth, A. (2006). Visual memory for natural scenes: Evidence from change detection and visual search. *Visual Cognition*, *14*(4–8), 781–807. https://doi.org/10.1080/13506280500193818

Holm, L., & Mantyla, T. (2007). Memory for scenes: Refixations reflect retrieval. *Memory and Cognition*, *35*(7), 1664–1674. https://doi.org/10.3758/BF03193500

Isola, P., Parikh, D., Torralba, A., & Oliva, A. (2011). Understanding the Intrinsic Memorability of Images. Retrieved from http://papers.nips.cc/paper/4451-understanding-the-intrinsic-memorability-of-images

Itti, L., & Koch, C. (2000). A saliency-based search mechanism for overt and covert shifts of visual attention. In *Vision Research*. https://doi.org/10.1016/S0042-6989(99)00163-7

Itti, L., & Koch, C. (2001). Computational modelling of visual attention. *Nature Reviews Neuroscience*, *2*(3), 194–203. https://doi.org/10.1038/35058500

Itti, L., Koch, C., & Niebur, E. (1998). A model of saliency-based visual attention for rapid scene analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. https://doi.org/10.1109/34.730558

Johnson, M. K. (1992). MEM: Mechanisms of Recollection. *Journal of Cognitive Neuroscience*, *4*(3), 268–280. https://doi.org/10.1162/jocn.1992.4.3.268

Kafkas, A., & Montaldi, D. (2011). Recognition memory strength is predicted by pupillary responses at encoding while fixation patterns distinguish recollection from familiarity. *The Quarterly Journal of Experimental Psychology*, *64*(10), 1971–1989. https://doi.org/10.1080/17470218.2011.588335

Kafkas, A., & Montaldi, D. (2012). Familiarity and recollection produce distinct eye movement, pupil and medial temporal lobe responses when memory strength is matched. *Neuropsychologia*, *50*(13), 3080–3093. https://doi.org/10.1016/j.neuropsychologia.2012.08.001

Kardan, O., Berman, M. G., Yourganov, G., Schmidt, J., & Henderson, J. M. (2015). Classifying mental states from eye movements during scene viewing. *Journal of Experimental Psychology. Human Perception and Performance*, *41*(6), 1502–1514. https://doi.org/10.1037/a0039673

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest Package: Tests in Linear Mixed Effects Models. *Journal of Statistical Software*, *82*(13), 1–19. https://doi.org/10.18637/jss.v082.i13

Loftus, G. R. (1972). Eye fixations and recognition memory for pictures. *Cognitive Psychology*, *3*(4), 525–551. https://doi.org/10.1016/0010-0285(72)90021-7

Maechler, M. (2015). Cluster:“Finding Groups in Data”: Cluster Analysis Extended Rousseeuw et al. R package version 2.0. 3.

Mancas, M., & Le Meur, O. (2013). Memorability of natural scenes: The role of attention. In *2013 IEEE International Conference on Image Processing* (pp. 196–200). IEEE. https://doi.org/10.1109/ICIP.2013.6738041

Mantyla, T., & Holm, L. (2006). Gaze control and recollective experience in face recognition.

Visual Cognition, *14*(3), 365–386. https://doi.org/10.1080/13506280500347992

Meister, M. L. R., & Buffalo, E. A. (2016). Getting directions from the hippocampus: The neural connection between looking and memory. *Neurobiology of Learning and Memory*, *134*, 135–144. https://doi.org/10.1016/j.nlm.2015.12.004

Mills, M., Hollingworth, A., Van der Stigchel, S., Hoffman, L., & Dodd, M. D. (2011). Examining the influence of task set on eye movements and fixations. *Journal of Vision*, *11*(8), 17–17. https://doi.org/10.1167/11.8.17

Montaldi, D., Spencer, T. J., Roberts, N., & Mayes, A. R. (2006). The neural system that mediates familiarity memory. *Hippocampus*, *16*(5), 504–520. https://doi.org/10.1002/hipo.20178

Paller, K. A., & Wagner, A. D. (2002). Observing the transformation of experience into memory. *Trends in Cognitive Sciences*, *6*(2), 93–102. https://doi.org/10.1016/S1364-6613(00)01845-3

Ramey, M. M., Yonelinas, A. P., & Henderson, J. M. (2019). Conscious and unconscious memory differentially impact attention: Eye movements, visual search, and recognition processes. *Cognition*, *185*, 71–82. https://doi.org/10.1016/J.COGNITION.2019.01.007

Roediger, H. L. (1990). Retention Without Remembering. *American Psychologist*, 1043–1056.

Rosenthal, R., & Rosnow, R. L. (1991). *Essentials of Behavioral Research: Methods and Data Analysis*. *McGraw-Hill Series in Psychology*.

Rousseeuw, P. J. (1987). *Silhouettes: a graphical aid to the interpretation and validation of cluster analysis*. *Journal of Computational and Applied Mathematics* (Vol. 20).

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

Schacter, D. L., Chiu, C.-Y. P., & Ochsner, K. N. (1993). *IMPLICIT MEMORY: A Selective Review*.

Schwedes, C., Scherer, D., & Wentura, D. (2019). Manipulating the depth of processing reveals the relevance of second eye fixations for recollection but not familiarity. *Psychological Research*, *1*, 3. https://doi.org/10.1007/s00426-019-01218-x

Schwedes, C., & Wentura, D. (2019). The relevance of the first two eye fixations for recognition memory processes. *Memory*. https://doi.org/10.1080/09658211.2019.1567789

Sharot, T., Davidson, M. L., Carson, M. M., & Phelps, E. A. (2008). Eye movements predict recollective experience. *PloS ONE*, *3*(8). https://doi.org/10.1371/journal.pone.0002884

Smith, C. N., Hopkins, R. O., & Squire, L. R. (2006). Experience-dependent eye movements, awareness, and hippocampus-dependent memory. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *26*(44), 11304–12. https://doi.org/10.1523/JNEUROSCI.3071-06.2006

Smith, C. N., & Squire, L. R. (2008). Experience-Dependent Eye Movements Reflect Hippocampus-Dependent (Aware) Memory. *Journal of Neuroscience*, *28*(48), 12825–12833. https://doi.org/10.1523/JNEUROSCI.4542-08.2008

Smith, C. N., & Squire, L. R. (2017). When eye movements express memory for old and new scenes in the absence of awareness and independent of hippocampus. *Learning & Memory*, *24*(2), 95–103. https://doi.org/10.1101/lm.043851.116

SR Research. (2010a). Experiment builder user’s manual. Mississauga, ON: SR Research Ltd.

SR Research. (2010b). EyeLink 1000 user’s manual, version 1.5.2. Mississauga, ON: SR Research Ltd.

Tatler, B. W., Hayhoe, M. M., Land, M. F., & Ballard, D. H. (2011). Eye guidance in natural vision: Reinterpreting salience. *Journal of Vision*, *11*(5), 5–5. https://doi.org/10.1167/11.5.5

Torralba, A., Oliva, A., Castelhano, M. S., & Henderson, J. M. (2006). Contextual guidance of eye movements and attention in real-world scenes: The role of global features in object search. *Psychological Review*, *113*(4), 766–786. https://doi.org/10.1037/0033-295X.113.4.766

Valuch, C., Becker, S. I., & Ansorge, U. (2013). Priming of fixations during recognition of natural scenes. Journal of Vision, 13(3), 1–22. https://doi.org/10.1167/13.3.3.doi

Van der Stigchel, S., Belopolsky, A. V., Peters, J. C., Wijnen, J. G., Meeter, M., & Theeuwes, J. (2009). The limits of top-down control of visual attention. *Acta Psychologica*, *132*(3), 201–212. https://doi.org/10.1016/j.actpsy.2009.07.001

Võ, M. L.-H., & Wolfe, J. M. (2015). The role of memory for visual search in scenes. *Annals of the New York Academy of Sciences*, *1339*(1), 72–81. https://doi.org/10.1111/nyas.12667

Whittlesea, B. W. A. (1993). *Illusions of Familiarity*. *Journal of Experimental Psychology: Learning, Memory, and Cognition* (Vol. 19).

Wynn, J. S., Bone, M. B., Dragan, M. C., Hoffman, K. L., Buchsbaum, B. R., & Ryan, J. D. (2016). Selective scanpath repetition during memory-guided visual search. *Visual Cognition*, *24*(1), 15–37. https://doi.org/10.1080/13506285.2016.1175531

Yonelinas, A. P. (2001). Components of episodic memory: The contribution of recollection and familiarity. In *Philosophical Transactions of the Royal Society B: Biological Sciences* (Vol. 356, pp. 1363–1374). https://doi.org/10.1098/rstb.2001.0939

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

Yonelinas, A. P., Otten, L. J., Shaw, K. N., & Rugg, M. D. (2005). Separating the Brain Regions Involved in Recollection and Familiarity in Recognition Memory. *Journal of Neuroscience*, *25*(11), 3002–3008.

Appendix

Post-hoc Analyses

Model comparison. To provide a direct test of the sensitivity of the presently developed measures of visual attention (i.e., dispersion and resampling) above and beyond previously used oculomotor measures (i.e., number of fixations and saccade amplitude) in predicting memory, we conducted a model comparison. Specifically, we determined whether adding dispersion and resampling predictors to a model regressing familiarity strength on number of fixations and saccade amplitude provided a better fit to the data. We found that adding dispersion and resampling produced a superior model in both experiments, at both encoding and retrieval, *p*s<.0001.

Experiment effects. To determine whether the differences between experiments (i.e., memory accuracy, and the use of a different set of scenes) led to differences in the relation between memory and the eye movement measures, we compiled both experiments into one data set and re-ran the analyses. Every significant effect obtained in either of the experiments alone replicated to *p*s<.007 in the combined data set: dispersion predicted familiarity strength and recollection during encoding and retrieval, and resampling predicted familiarity strength and recollection. When a covariate of the experiment to which each data point belonged was included in the model, all of the effects held; this covariate of experiment was not significant in any contrast, *p*s>.07. There were also no interactions between experiment of origin and memory in predicting the eye movement measures, *p*s>.05.

Furthermore, one possible concern is the addition of outdoor scenes in Experiment 2, which were included to increase the distinctiveness of the stimuli and therefore the proportion of high-confidence memory responses, whereas only indoor scenes were used in Experiment 1. In analyses of data from Experiment 2, we found that the indoor versus outdoor status of a scene did not interact with memory in predicting any of the eye movement effects. Together, these analyses suggest that the results were robust across different scenes and accuracy rates.

Collapsing all misses. It is possible that the lack of effect of unconscious memory on dispersion could be related to the strict criterion we used for the unconscious memory contrast (i.e., "sure new" scenes, that is, high-confidence misses). Therefore, to ensure that this was not responsible for the null effects, we collapsed "sure new" and "maybe new" responses into a single "new" response and re-ran the retrieval dispersion analysis. There was still no effect in Experiment 1, $t(921)= 0.54$, $p=0.59$, $d=0.04$, or Experiment 2, $t(2300)= 1.56$, $p=0.12$, $d=0.07$. Thus, even with approximately double the number of trials included in the analysis and a more lax definition of misses, there was no effect of unconscious memory on dispersion.

Correlations between eye movement measures. Both number of fixations and saccade amplitude have previously been interpreted at times to reflect spatial distribution and clustering of eye movements during viewing, but neither measure takes spatial information into account. To determine if these measures may serve as a proxy for spatial distribution, we assessed their correlation with each other and the presently developed measure of dispersion using the combined data from both experiments (Table A1). Interestingly, despite the fact that both measures have been interpreted to represent spatial distribution of eye movements, saccade amplitude and number of fixations were only weakly associated, $r = .09$, $p <.001$. However, both measures were moderately correlated with dispersion, *r*s > .41, *p*s <.001 (Table A1). Moreover, when number of fixations and saccade amplitude were submitted to a principal component analysis, the first principal component exhibited a stronger correlation with dispersion than either measure did individually, $r = .87$. This suggests that both number of fixations and saccade amplitude do contain non-overlapping subsets of information about spatial distribution, such that they converge towards describing spatial distribution when combined, but they do not provide a complete picture of how widely viewing is distributed across a stimulus.

Table A1

Pearson Correlations Between Measures Related to Dispersion, Across Both Experiments

	# fixations	Sac. Amplitude	# regions	Dispersion
# fixations	1	.09	.71	.45
Sac. Amplitude	.09	1	.33	.41
# regions	.71	.33	1	.50
Dispersion	.45	.41	.50	1

Note: All correlations were significant to *p* <.0001. Sac. Amplitude = saccade amplitude.

Additional Analyses. To increase the interface of the present findings with the memory and visual attention literature in general, we ran supplemental analyses. First, to confirm the effect of memory on dispersion, we examined whether dispersion significantly differed between study and test as well as between old and new scenes at test. For old scenes, dispersion was lower at test than at study in Experiment 1, *t*(6366)= -3.42, *p*<.001, *d*= -0.09, and Experiment 2, *t*(13218)= -21.44, *p*<.0001, *d*= -0.37. Similarly, at test, dispersion was significantly lower for old than new scenes in Experiment 1, *t*(4065)= -4.07, *p*<.0001, *d*= -0.13, and Experiment 2, *t*(8654)= -8.6, *p*<.0001, *d*= -0.18. Second, average fixation duration was examined. During encoding, fixation duration was related to familiarity in Experiment 1, β = -0.082, *t*(2148)= -3.92, *p*<.0001, and Experiment 2, β = -0.088, *t*(3633)= -5.22, *p*<.0001. Fixation duration during encoding was also related to recollection, albeit weakly, in Experiment 1, *t*(1743)= -2.05, *p*= .041, *d*= -0.1, and Experiment 2, *t*(4761)= -2.15, *p*= .032, *d*= -0.06. At retrieval, fixation duration was not related to familiarity in Experiment 1 or Experiment 2, *p*s>.13. Fixation duration during retrieval was not related to recollection in Experiment 1, *t*(1730)= 1.34, *p*= .18, *d*= 0.06, but it was in Experiment 2, *t*(4771)= 3.54, *p*<.001, *d*= 0.1. There were no unconscious effects on fixation duration, *p*s>.44.

Chapter 3

Why do we retrace our visual steps?

Semantic and episodic memory in gaze reinstatement

Michelle M. Ramey[a,b,c*], Andrew P. Yonelinas[a,b], John M. Henderson[a,c]

[a] Department of Psychology, University of California, Davis, CA, USA.

[b] Center for Neuroscience, University of California, Davis, CA, USA.

[c] Center for Mind and Brain, University of California, Davis, CA, USA.

Published in *Learning and Memory* (2020)

Author Note:

Funding: This work was supported by the National Eye Institute of the National Institutes of Health under Award Numbers R01EY025999 and R01EY027792. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Abstract

When we look at repeated scenes, we tend to visit similar regions each time—a phenomenon known as *resampling*. Resampling has long been attributed to episodic memory, but the relationship between resampling and episodic memory has recently been found to be less consistent than assumed. A possibility that has yet to be fully considered is that factors unrelated to episodic memory may generate resampling: for example, other factors such as semantic memory and visual salience that are consistently present each time an image is viewed and are independent of specific prior viewing instances. We addressed this possibility by tracking participants' eyes during scene viewing to examine how semantic memory, indexed by the semantic informativeness of scene regions (i.e., *meaning*), is involved in resampling. We found that viewing more meaningful regions predicted resampling, as did episodic familiarity strength. Furthermore, we found that meaning interacted with familiarity strength to predict resampling. Specifically, the effect of meaning on resampling was attenuated in the presence of strong episodic memory, and vice versa. These results suggest that episodic and semantic memory are each involved in resampling behavior and are in competition rather than synergistically increasing resampling. More generally, this suggests that episodic and semantic memory may compete to guide attention.

Keywords: Memory; recognition; semantic memory; attention; eye movements

Why do we retrace our visual steps?

Semantic and episodic memory in gaze reinstatement

Most people have heard of retracing one's steps to find something lost, but this idea goes beyond a useful adage: Retracing one's steps may be an important part of episodic memory. For example, returning to the location in which one learned information enhances the ability to retrieve that information by reinstating the encoding context (Godden and Baddeley 1975; Smith 1979). There are also theories proposing that this effect extends to where we look—such that reinstating gaze by viewing similar regions across study and test of an image (i.e., resampling) improves episodic memory for that image—based on observations of consistent eye movement patterns between successive viewings of an image (Noton and Stark 1971). Though controversial (Henderson 2003), such theories have had a longstanding influence on the literature in visual attention and memory (Wynn, Shen, & Ryan, 2019). However, research has since indicated that the relationship of resampling with episodic memory is more tenuous than had been assumed, and the causality of the relationship has remained elusive (e.g., Damiano & Walther, 2019; Foulsham & Kingstone, 2013; Holm & Mantyla, 2007; Locher & Nodine, 1974; Valuch, Becker, & Ansorge, 2013). Despite this tenuous relationship and the rapidly growing interest in resampling behavior (e.g., Wynn et al. 2019), attempts to explain what gives rise to resampling have continued to focus primarily on episodic memory as the driving factor (Foulsham and Kingstone 2013; Wynn et al. 2016, 2018; Holm and Mantyla 2007; Damiano and Walther 2019; Valuch et al. 2013), even in studies finding no relation between recognition performance and resampling (e.g., Humphrey and Underwood 2010; Locher & Nodine, 1974). Because of this longstanding focus on episodic memory in the resampling literature, it is not well understood how other cognitive or visual factors may contribute to resampling behavior, and how such factors may modulate the relationship between resampling and episodic memory.

The notion that resampling image regions between successive viewings is a uniquely episodic-memory-related phenomenon began with Noton and Stark's (1971) scanpath theory. They proposed that the path traveled by the eyes during encoding was stored alongside memory representations for the visual information, and that repeating the scanpath upon subsequent viewings of the image facilitated memory retrieval. This theory was based upon observations that participants tend to produce similar scanpaths between repeated viewings of a given image, and it was assumed that this resampling behavior was a uniquely memory-related phenomenon—without any direct test of whether the repeated scanpaths related to episodic memory. In fact, the first direct tests found no relationship between resampling and memory accuracy (e.g., Humphrey and Underwood 2010; Locher and Nodine 1974). Only recently has evidence surfaced for a relationship, albeit weak, between resampling and recognition memory performance (Valuch et al. 2013; Holm and Mantyla 2007; Mantyla and Holm 2006; Foulsham and Kingstone 2013; Wynn et al. 2016; Damiano and Walther 2015, 2019; Ramey, Henderson, & Yonelinas, 2020). Furthermore, most of the evidence thus far has been correlational. The few causal studies that have been done have found consistent evidence for an influence of episodic memory on resampling behavior, but often find no significant influence of resampling on episodic memory (Holm and Mantyla 2007; Foulsham and Kingstone 2013; Valuch et al. 2013; Damiano and Walther 2015)—despite the assumptions of scanpath theory and other similar theories that resampling improves episodic memory through an iterative, bidirectional process (Noton and Stark 1971).

Thus far, the evidence suggests that episodic memory and resampling are related, and that this is primarily driven by stronger memory causally increasing the extent to which similar image regions are viewed. However, a substantial amount of resampling behavior occurs that is not explained by variations in recognition memory strength (Ramey et al. 2020), which suggests that there may be additional driving forces behind resampling. One possibility is that fixating similar regions between repeated viewings of an image could simply reflect consistent guidance by factors that are present irrespective of whether the image is remembered, such as other known sources of influence on eye movement behavior (e.g., visual salience and general world knowledge; see Henderson, 2003). For example, viewers tend to look at semantically informative scene regions (e.g., regions containing objects) more than uninformative regions (e.g., empty regions; Henderson & Hayes, 2017). If they tend to do this for both initial and repeated viewings, then what might appear to be memory-based resampling could instead simply be due to selection of the same regions independently across viewings. Despite the intuitive appeal of and indirect support for such possibilities, they have yet to be directly investigated. The present study is thus aimed at determining whether additional factors, primarily the semantic informativeness of scene regions, are involved in resampling behavior, and whether such factors may influence the extent to which episodic memory predicts resampling.

A direct measure of the spatial distribution of semantic information in scenes was recently developed by constructing *meaning maps* based on participants' ratings of different parts of each scene (Henderson and Hayes 2017). Meaning maps are able to capture the spatial distribution of potentially useful information in a scene, such as objects and people, and thus might be expected to provide a good estimate of where people tend to look. Indeed, quantification of the amount of meaning contained in image patches has been found to quite accurately predict the spatial distribution of attention (Henderson and Hayes 2017, 2018). As described above, meaning may naturally lead to resampling behavior because the distribution of meaningful regions in scenes is unchanged between viewings, and one would thus expect attention to be driven to meaningful information similarly between viewings.

Given that meaning indexes the semantic informativeness of scene regions, it may serve as a measure of semantic memory. That is, objects are only informative insofar as we have learned that they are informative, through a lifetime of accruing semantic knowledge about the world (Saffran and Schwartz 1994; Tulving 1986). Therefore, the fact that people tend to direct their attention towards meaningful regions suggests that semantic memory consistently influences attention during naturalistic viewing. This possibility may prove particularly relevant for the resampling literature—given its longstanding focus on episodic memory—as there is emerging evidence from the visual search literature that semantic and episodic memory can interact to influence attention. For example, in addition to studies showing that episodic and semantic memory are each able to guide viewing behavior generally (Henderson, 2003; Neider & Zelinsky, 2006; Ryan & Shen, 2020; Wynn et al., 2020), one study suggests that decreasing the extent to which semantic memory is available to guide search leads to an increase in reliance on episodic memory (Võ and Wolfe 2013). This indicates that semantic and episodic memory may trade off or compete in their guidance of attention during search. It is therefore possible that the semantic meaning of viewed scene regions may influence the relationship between episodic memory and resampling behavior as well.

In sum, the evidence thus far indicates that episodic memory is indeed related to resampling behavior, such that stronger episodic memory increases the extent to which regions are resampled (Holm and Mantyla 2007; Foulsham and Kingstone 2013; Valuch et al. 2013;

Damiano and Walther 2015). Despite this, much of resampling behavior remains unexplained by episodic memory; the resampling phenomenon is being hotly investigated in the memory literature, but it is not yet well understood why people tend to revisit regions between study and test. As mentioned above, however, there is also indirect support for the possibility that semantic memory may be able to produce resampling behavior (Henderson & Hayes, 2017), and that it may interact with episodic memory to do so (Võ & Wolfe, 2013; Wynn et al., 2020).

Current Research

To address these possibilities, we examined how resampling behavior is related to the semantic informativeness of viewed scene regions, and whether directing attention towards semantically informative regions may modulate how episodic memory is involved in resampling. Furthermore, we examined resampling on a more fine-grained level than in prior work by developing a new fixation-by-fixation measure for assessing the extent to which each retrieval fixation was near regions that had been visited during encoding. This new measure (i.e., *refixation distance*) also allows fixation-by-fixation trends in how spatial resampling is related to memory to be examined for the first time; that is, it allows for assessment of how resampling varies over the course of a trial. We assessed recognition memory using a confidence-based memory scale, to allow for a sensitive assessment of memory strength (Ramey, Yonelinas, & Henderson, 2019), rather than previously used dichotomous old/new judgments.

In the present experiment, participants viewed a series of real-world scenes while their eye movements were tracked. During an initial study phase, participants viewed scenes in two encoding tasks. In one encoding task, participants were asked to memorize each scene, whereas in the other, they were asked to judge each scene for its aesthetic appeal. Two encoding tasks were included to test the generalizability of any effects obtained as well as to verify that the effects were not limited to conditions in which participants intentionally encoded the scenes. During a subsequent test phase, participants viewed the same scenes that they had viewed during the study phase (i.e., old scenes) along with randomly intermixed new scenes, and were asked to provide a recognition judgment for each scene. Recognition memory awareness was measured by asking participants to rate their memory confidence for each scene on a 6-point scale during the recognition judgment (Yonelinas 2002). Participants were told that if they could consciously recollect some qualitative aspect of the initial learning event, such as what they thought about when the scene was encountered earlier, they should respond "Recollect old (6);" otherwise, they rated their memory confidence by responding "I'm sure it's old (5)," "Maybe it's old (4)," "I don't know (3)," "Maybe it's new (2)," or "I'm sure it's new (1)." In a prior study of scene memory, we found that trial-by-trial resampling was consistently related to familiarity strength, but not recollection (Ramey et al. 2020); therefore, we focused the present analyses on the continuous gradient of familiarity-based memory responses from "sure new" to "sure old." However, for the sake of completeness, we also verified that the effects held when using all responses (see Appendix), and we present the recollect responses in our data figures.

Semantic informativeness was quantified using meaning maps (Hayes & Henderson 2017), which capture the spatial distribution of semantic information across a scene. Attention to meaning was used as an index of attentional guidance by semantic memory and was computed on a fixation-wise basis by determining the average amount of meaning contained in the region immediately surrounding each fixation. Furthermore, to ensure that attention to meaningful regions was not potentially confounded by bottom-up visual saliency, we also ran analyses that controlled for saliency and examined the effects of saliency on resampling (see Appendix).

Because the meaning of scene regions has been shown to be a potent driver of attention, and there is no evidence to our knowledge suggesting this would not be the case across multiple viewings, we hypothesized that fixating highly meaningful regions would be a strong predictor of resampling. The potential for interaction between semantic meaning and episodic memory is less clear, but a variety of outcomes would be of theoretical relevance. First, it is possible that semantic meaning and episodic memory may have a synergistic relationship in predicting resampling. This could be the case if resampling improves memory, as some theories have proposed (Wynn et al. 2019), and if attention to meaning increases resampling. A second possibility is that meaning and episodic memory may compete with each other to guide attention, such that a stronger influence of one on any given fixation might lessen the influence of the other. This would fit with the emerging evidence of potential competitive interactions between semantic and episodic memory (Võ and Wolfe 2013).

Results

Preliminary Analyses

Recognition memory accuracy. The percentage of scenes that received a recognition confidence response of “recollect,” “sure old,” “maybe old,” “don’t know,” “maybe new,” and “sure new,” respectively, were 46%, 25%, 11%, 8%, 6%, and 4% for old scenes, and 2%, 3%, 8%, 13%, 27%, and 47% for new scenes (Figure 2a). These results suggest that participants were able to discriminate between old and new scenes and used the full range of response options.

Study task. We included two study tasks (i.e., memorization and aesthetic judgment) to ensure that any effects obtained were robust to encoding conditions. To determine if study task affected resampling behavior, we ran three models: 1) regressing refixation distance on study task, 2) regressing refixation distance on the interaction between study task and memory strength, and 3) regressing refixation distance on the interaction between study task and meaning score. There were no significant effects of study task (*p*s >.08). Subsequent analyses are thus collapsed across tasks.

Primary Analyses

For a schematic of how episodic memory and semantic meaning may be able to drive resampling behavior via individual fixations, see Figure 1. Note that all of the analyses below use data from old scenes in the test phase, because it was not possible to directly assess resampling in new scenes.

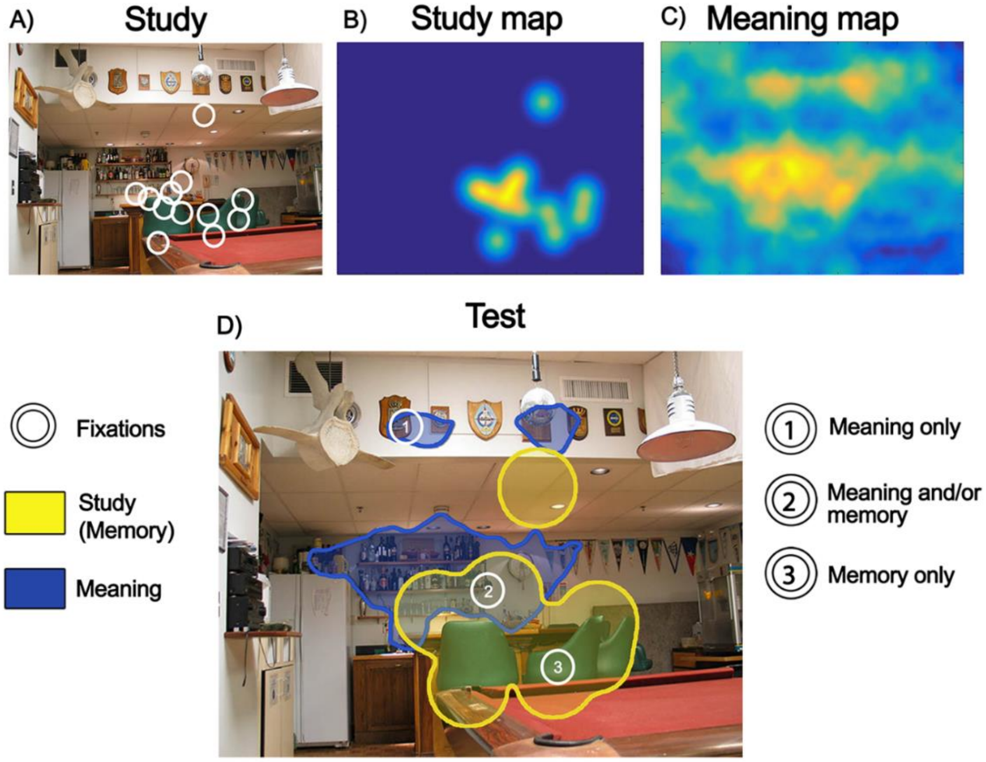


Figure 1. Conceptual framework using example fixations made during study and test of a scene. A) Fixations (rings) made during study of a scene. B) A smoothed heatmap of the study fixation locations from the same scene. C) The meaning map of the same scene. D) A combination of the scene and the brightest portions of the heatmaps in b-c), to illustrate how semantic meaning and episodic memory may drive resampling. Fixation 1 represents a non-resampled fixation that was likely driven by meaning (blue). Fixation 2 is an example of resampling (i.e., is within a yellow region) that could be driven by memory and/or meaning given that they are overlapping. Fixation 3 is an example of resampling that was likely not driven by meaning. We hypothesized that an increase in memory strength would lead to an increase in fixations like fixations 2 and 3, whereas attention to meaning would result in fixations like 1 and 2.

Episodic memory strength. To examine how resampling behavior related to recognition memory strength, we determined whether refixation distance varied across the linear gradient of "sure new" to "sure old" responses (i.e., familiarity-based memory strength). Memory strength predicted significantly decreased refixation distance, β= -.14, *p*<.0001, indicating that stronger memory was related to increased resampling behavior (Figure 2b). Note that memory strength was a trial-level measure, whereas refixation distance was a fixation-by-fixation measure; we used a nested linear mixed effects model to account for this difference, but found that the effect remained when refixation distance was aggregated by trial as well, β= -.29, *p*<.0001 (and by subject; see Appendix).

Semantic meaning. To determine whether the semantic informativeness of viewed scene regions predicted resampling behavior, we regressed the refixation distance of each fixation on the meaning score of each fixation in the test phase. Meaning scores were negatively related to refixation distance, β= -.28, *p*<.0001, such that the tendency to resample a region was associated with the meaning of that region (Figure 2c). This indicates that semantic memory may give rise to resampling behavior by guiding attention towards meaningful regions consistently across viewings.

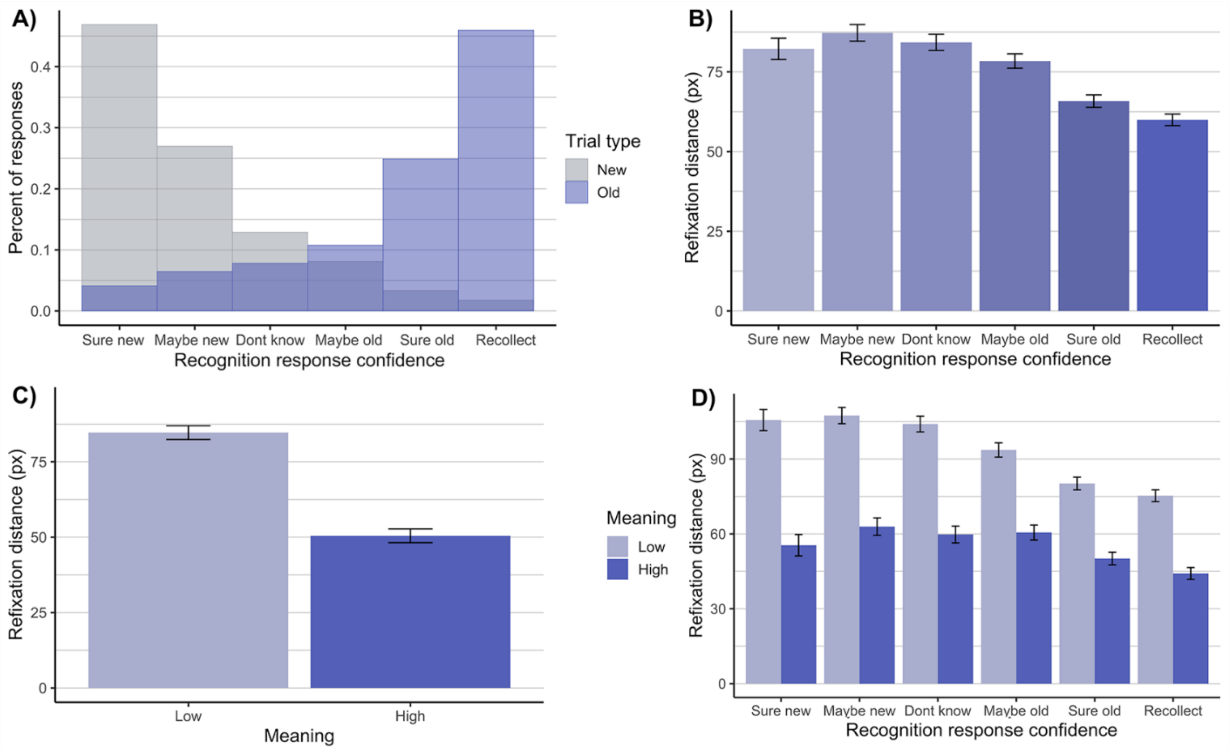


Figure 2. Results of primary resampling analyses. Note that lower refixation distance reflects more resampling behavior. A) Histogram of the proportions of recognition memory responses made for old and new scenes. B) Refixation distance by memory response. C) Refixation distance by attention to meaning (i.e., meaning score). D) Interaction between meaning score and memory response in predicting refixation distance. For b-d), least-squares means derived from the linear mixed effects models used in the analyses are plotted, and the error bars represent the standard error of these estimated means from the model. For c-d), meaning was dichotomized to facilitate visualization, but all analyses were done with continuous data.

Interaction between semantic meaning and episodic memory. The evidence thus far indicates that episodic memory and semantic meaning are each uniquely involved in resampling behavior: Memory strength was not related to the meaning of viewed regions (Eq. A6), and each predicted unique variance in resampling behavior (see Appendix). To confirm that each variable had a main effect on resampling when the other was controlled for, we ran a model predicting refixation distance from both meaning and memory strength. As expected, memory strength, β= -.11, *p*<.0001, and meaning, β= -.27, *p*<.0001, each predicted resampling when the other was held constant statistically.

Given that both factors appear to be simultaneously involved in resampling, we sought to determine whether semantic meaning may modulate the relationship between episodic memory and resampling based on preliminary evidence for interactions between episodic and semantic memory (Võ and Wolfe 2013). To do this, we ran a model regressing refixation distance on the interaction between meaning score and memory strength (Eq. A7). As predicted, meaning score and memory strength interacted to predict refixation distance, β= .06, *p*<.0001, such that an increase in meaning score reduced the strength of the relation between memory strength and resampling, and vice versa (i.e., an increase in memory strength reduced the association between meaning and resampling; Figure 2d). These results suggest that stronger semantic guidance attenuates the extent to which episodic memory strength predicts resampling, and/or that strong episodic memory may weaken the relationship between semantic memory and resampling.

Temporal Analyses

Because refixation distance is a fixation-by-fixation measure of resampling, it allows for examination of how the effects of episodic memory and semantic meaning on resampling may change over the course of a trial. To determine how the effects observed thus far changed over fixations, we regressed refixation distance on the interaction between the variable of interest and the ordinal fixation number in a trial.

Episodic memory strength. The relationship between recognition memory response and refixation distance did not change over fixations in a linear fashion, β= -.003, *p*=.58. However, examination of the plot (Figure 3a, Supplemental Figure 4) revealed what appeared to be a quadratic relationship, so we conducted an exploratory analysis of this possibility. Including a quadratic interaction term in the model, we found that the relationship between memory strength and refixation distance was significantly more pronounced towards the middle of the trial, β

=.03, *p*<.001. That is, the effect of memory strength on resampling was weakest at the beginning and end of each trial, and was strongest midway through each trial.

Semantic meaning. Regressing refixation distance on the interaction between meaning score and ordinal fixation number revealed a significant linear interaction, β = -.04, *p*<.0001, such that the relationship between meaning and refixation distance grew stronger over the course of the trial (Figure 3b). Unlike memory strength, there was no significant quadratic effect, β = .005, *p*=.53.

Interaction between semantic meaning and episodic memory. We next examined how the interaction between meaning and memory strength observed above might change over time, and found that it does not vary systematically over the course of the trial, *p*s>.39 (Supplemental Figure 3; see Appendix).

Taken together, the temporal analyses indicate that memory strength exhibited little effect on resampling early on in viewing, had a strong influence mid-trial, and was attenuated towards the end of the trial. In contrast, meaning demonstrated an effect early in the trial that consistently increased over the course of viewing. This suggests that semantic meaning and episodic memory might have unique time courses in how they relate to resampling over the course of viewing.

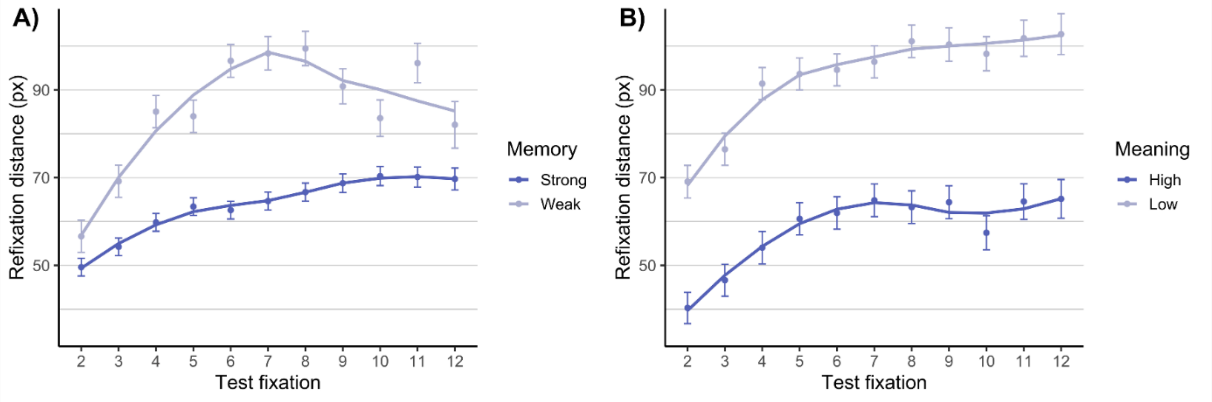


Figure 3. Results of the temporal analyses. A) Refixation distance over fixations by memory strength; “strong” memory included responses of “recollect” and “sure old,” and “weak” memory included responses of “sure new” and “maybe new.” (All responses are shown in

Supplemental Figure 4.) B) Refixation distance over fixations by the meaning score of each fixation. In each plot, the x-axis is the ordinal fixation number in a test phase trial. Ninety percent of the data were included at a cutoff of 12 fixations; the plot was thus truncated at 12 fixations to reduce noise from the small number of trials containing more than 12 fixations. However, analyses included all data. Least-squares means derived from the linear mixed effects models used in the analyses are plotted, and the error bars represent the standard error of these estimated means from the model. The lines were generated using a locally weighted smoothing function, which plots local regressions to aid the eye in seeing trends. The data in the plots were dichotomized to facilitate visualization, but all analyses were done with continuous data.

Additional Analyses. We ran a series of additional analyses that are presented in the Appendix. First, to probe the robustness of the present effects to potentially confounding variables, we re-ran analyses controlling for variables such as the number of fixations per trial, potential center bias of meaning maps, and the inclusion or exclusion of recollect responses—none of which altered the pattern of results. We also created cross-subject and cross-image refixation distance baselines by randomly pairing subject and image data, and found that resampling behavior is indeed driven by both subject and image-level idiosyncrasies such as episodic memory and semantic meaning of regions, respectively. That is, the observed refixation distance values were significantly lower than the refixation distance values obtained by randomly pairing images, or by randomly pairing subjects. Furthermore, the analysis suggests that image properties—such as semantic meaning—may be responsible for the majority of resampling behavior (Supplemental Figure 1b).

We then examined whether the bottom-up visual salience of scene regions related to resampling, and found that salience largely followed the same pattern of results as meaning. Furthermore, both meaning and salience accounted for unique variance in resampling behavior; thus, the relationship between meaning and resampling held when salience was controlled, indicating that the meaning effects observed above were driven by semantic information and not by bottom-up visual information. We also ran a combined model predicting refixation distance from memory strength, meaning, salience, center bias, and the interaction between meaning and memory strength, and found that every variable significantly predicted unique variance and improved the model fit.

Discussion

In the present study, we examined how episodic and semantic memory predicted resampling behavior, operationalized as the extent to which people revisited scene regions between study and test. We tracked participants' eye movements during encoding and retrieval of scenes, and participants provided confidence-based recognition judgments for each scene during retrieval. Resampling—as well as guidance by semantic memory, indexed by attention to meaningful regions (i.e., *meaning*)—were assessed on a fixation-by-fixation basis, allowing for fine-grained analysis of temporal trends. We found that episodic memory strength and semantic meaning each predicted resampling behavior, such that stronger recognition memory and increased viewing of meaningful regions during retrieval were both related to increased resampling. Importantly, memory strength was not related to meaning, and these factors each predicted unique variance in resampling behavior. Furthermore, episodic memory strength interacted with semantic meaning to predict resampling, such that stronger semantic guidance weakened the relationship between episodic memory and resampling, and vice versa. Moreover, these effects were not due to the numbers of fixations made, center bias of viewing, nor overall similarities in viewing patterns between subjects or scenes. We also found that these effects were robust to encoding conditions (i.e., memorization and aesthetic judgment), and that the effects of meaning were not driven by visual salience; in fact, meaning and salience each predicted unique variance in resampling. Lastly, episodic memory and meaning had different patterns of results in how they were related to refixation distance over the course of fixations. Specifically, whereas meaning influenced viewing from the first fixation onward—and its influence continued to strengthen over the course of a trial—memory strength had little effect early on, and its effect peaked midway through the viewing period. These results suggest that episodic and semantic memory might have fundamentally different time courses in their influence on attention.

In contrast to theories assuming that resampling behavior is uniquely related to episodic memory (Noton and Stark 1971), the present results indicate that there is a robust role for information present in the image itself—such as semantic informativeness and visual salience—in driving resampling as well. In fact, when considered in the same model, meaning was a stronger predictor of resampling than was memory strength. Additionally, computing resampling with randomly shuffled pairings of trials revealed that image content such as the meaning of regions may be a stronger contributor to resampling behavior than subject-level factors such as episodic memory. Whereas there is debate surrounding the mechanisms underlying the relationship between episodic memory and resampling (Henderson 2003; Noton and Stark 1971), the mechanisms driving the involvement of semantic memory in resampling may be more straightforward. For example, because semantic memory is known to guide attention (Henderson and Hayes 2017), it is likely that the meaning of regions predicts resampling behavior simply by guiding attention consistently with each viewing. This potential mechanism was supported by follow-up analyses indicating that increased attention to meaning at study strengthened the relationship between resampling and meaning at test (see Appendix). Furthermore, the association between semantic memory and resampling may reflect consistent attention to the relationships between semantically relevant scene elements, paralleling theories of resampling in relational episodic memory (Wynn et al., 2019). In addition to the need to incorporate a role for semantic and image factors in theories of resampling, the present results point to the need to consider potential modulatory roles of these factors in the relationship between episodic memory and resampling. That is, the finding that more attention to meaning weakens the extent to which episodic memory strength predicts resampling, and vice versa, suggests that strong guidance by semantic memory may reduce the extent to which episodic memory is able to guide attention.

The apparent competition between semantic meaning and episodic memory observed in the present study also has potentially important implications for theories of attention. Specifically, many theories of attention have focused on competition between bottom-up or perceptual sources of guidance, such as image salience, and top-down cognitive factors in guiding eye movements (Tatler et al. 2011; Van der Stigchel et al. 2009). The present results, however, also point to the possibility that different top-down factors (i.e., semantic and episodic memory) may compete with each other to guide naturalistic viewing—a possibility that has been less well explored, particularly with respect to episodic memory. Harnessing meaning maps as a new method of indexing the distribution of semantic information, combined with the use of resampling as an index of attentional deployment—rather than solely as an index of episodic memory as has been done in prior work—provides a unique new window through which we can observe competition between semantic and episodic memory during naturalistic viewing for what is, to our knowledge, the first time. Using these measures, the current study suggests that fixations may be the result of a conflict between semantic and episodic memory, among other factors, to determine where attention is deployed. For example, when strong semantic guidance is present, episodic memory is less likely to "win" on any given fixation, and semantic memory would thus emerge as the stronger driver of resampling. However, further investigation of the potential interplay between episodic and semantic memory in driving attention is warranted before causal conclusions can be drawn.

In addition to implications for theories of both memory and attention, the present findings provide a potential new lens through which to view prior investigations of resampling behavior. The presently identified importance of both visual and semantic image content in resampling behavior indicates that prior findings of a relationship between episodic memory and resampling could have, in part, been driven by image content. Specifically, in the present study, we incorporated a random effect of image in all analyses—including the null relation between recognition memory strength and attention to meaning—to ensure that we were not including potentially confounding image effects such as overall differences in meaning, salience, or memorability. The majority of prior studies, however, have not controlled for such image effects. This is a particularly important consideration because of findings that images that are more memorable also tend to elicit more similar viewing patterns between subjects (Mancas and Le Meur 2013). When image effects are not controlled for, this effect could emerge as an apparent within-subjects relation between resampling and episodic memory, when in fact it is a result of certain images leading to more stereotyped scan patterns—even between subjects—and better memory. Therefore, accounting for potential confounds of image properties may be particularly important for future investigations of resampling and memory.

Taken together, the present findings indicate that resampling behavior reflects cognitive sources of guidance besides episodic memory, and that these factors may influence the relationship between episodic memory and resampling. Future investigations aimed at uncovering other such factors that guide resampling, and how they might modulate its relationship with episodic memory, may prove fruitful. In particular, these results highlight the complex interplay of cognitive and visual factors that orchestrate how we guide our attention: it is rarely, if ever, just one factor at play. Our knowledge of the world, the task at hand, our memories, the current visual input, and likely myriad other influences are all resolved within a few hundred milliseconds to produce each movement of the eyes.

Materials and Methods

Participants. Forty-five undergraduates from the University of California, Davis completed the experiment for course credit. The sample size was selected to provide more than 98% power to detect the weakest effect of subjectively reported memory on eye movements obtained in a prior study (Ramey et al., 2019). The quality of each participant's eyetracking data was assessed by computing the mean percent signal across all trials to determine whether there was excessive track loss due to blinks or calibration loss. All participants had greater than the preselected criterion of 75% signal (*M* = 94.7%), (Henderson & Hayes, 2017), such that they lost less than 25% signal; all participants were thus retained for analysis.[7]

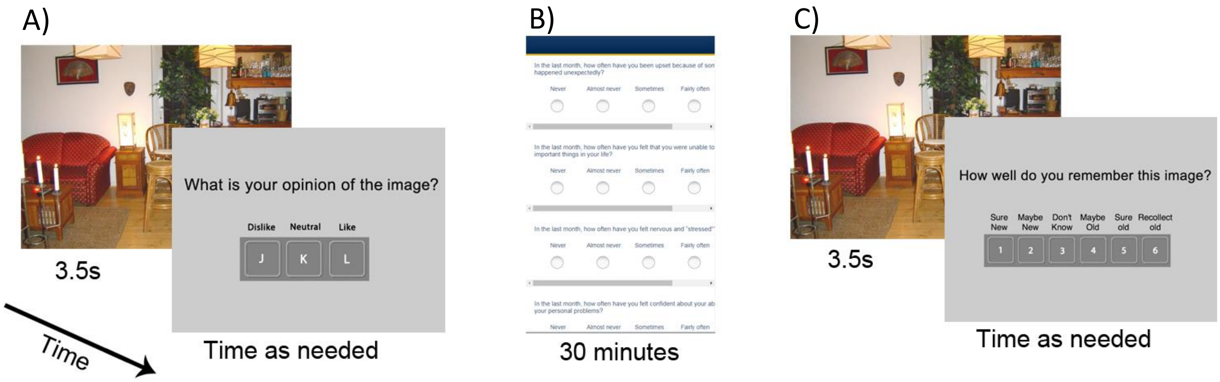
Stimuli. Stimuli were 200 photographs of real-world scenes. All scenes were presented in color at 1024x768 pixels subtending a visual angle of approximately 25°x19° at presentation. Of these 200 scenes, 150 were presented at study and test, and 50 were presented only at test. Eighty out of the 200 scenes had been run through the meaning mapping procedure (from Henderson & Hayes, 2017) and were used in analyses. Stimulus presentation was counterbalanced, such that each scene appeared in different conditions (i.e., in one of the two study tasks, or as a new lure during test; see procedure) for different participants, to mitigate stimulus effects.

Apparatus. Participants' eye movements were recorded using an SR Research EyeLink 1000+ tower mount eyetracker, sampling at 1000hz. A forehead and chin rest were used to reduce head movements, and eye movements were recorded from one eye though viewing was

---

[7] This data set is the same as that used in Experiment 2 of Ramey, Henderson, & Yonelinas (2020), which focused on a separate set of questions pertaining to episodic memory processes, but contained analyses of trial-level resampling and memory processes. Ramey et al. (2020) did not include any meaning or salience data, or fixation-level measures.

binocular. Stimuli were displayed on a monitor 85cm from the eyetracker, and the experiment was controlled with SR Research Experiment Builder software (SR Research, 2010a).



In the last month, how often have you been upset because of son happened unexpectedly?

In the last month, how often have you felt that you were unable to important things in your life?

In the last month, how often have you felt nervous and "stressed"

In the last month, how often have you felt confident about your ab your personal problems?

Figure 4. Illustration of the procedure. A) Study phase. Half of the scenes were presented in an aesthetic judgment task (i.e., participants were instructed to judge the image aesthetically and rate it as “dislike,” “neutral,” or “like”), whereas the other half were presented in a memorization task (i.e., participants were instructed to memorize the image and rate it as “not memorable,” “neutral”, or “memorable”). B) Delay between study and test, during which participants completed unrelated questionnaires. C) Test phase in which participants rated their recognition confidence.

Procedure. The experiment lasted 1.5 hours and consisted of a study phase followed by a filled 30min delay, as well as a subsequent test phase (see Figure 4). Eye movements were recorded throughout the study and test phases. In both phases, each trial (i.e., each scene presentation) was preceded by a central fixation cross. Participants were given breaks every 50 trials and between phases, and the eyetracker was recalibrated after each of these breaks.

Study phase. During the study phase, participants were presented with 150 unique scenes split into two task blocks: an aesthetic judgment task and a memorization task. These tasks were selected to ensure that any effects obtained were not a product of a given task, but rather generalized across tasks (as prior work has shown that eye movements vary systematically between tasks; Castelhano, Mack, & Henderson, 2009; Kardan, Berman, Yourganov, Schmidt, & Henderson, 2015; Mills, Hollingworth, Van der Stigchel, Hoffman, & Dodd, 2011). The order of the tasks was counterbalanced such that half of the participants completed the aesthetic judgment task first, whereas the other half completed the memorization task first. In each task, 75 scenes were presented for 3.5s each, allowing for an average of 12 fixations per trial. Each task was preceded by two practice trials to familiarize participants with the procedure.

In the aesthetic judgment task of the study phase, participants were asked to rate each scene based on how aesthetically pleasing they found it to be. Each trial consisted of the scene presentation, followed by a grey response screen containing the prompt “What is your opinion of the photo?” as well as the key mappings for each response option. Responses were made on the keyboard, had no time limit, and consisted of “dislike,” “neutral,” and “like;” the response data were not used.

The memorization task of the study phase followed the same general procedure, but participants were instead asked to memorize the scenes. After each scene, they were asked to rate how memorable they found the scene to be. Participants were asked to give this response to ensure that the sequence of events in the memorization task was analogous to the aesthetic judgment task. Responses included “not memorable,” “neutral,” and “memorable;” again, the response data were not used.

Delay. Between the study and test phases, participants were moved to a computer in a different room to complete a 30min distractor task that included questionnaires (e.g., personality scales) that were not related to the present study.

Test phase. In the test phase, participants were presented with a series of scenes and asked to rate their recognition memory for each scene. The test phase consisted of 200 trials: 150 old scenes, which had been presented in the study phase, and 50 randomly intermixed new scenes, which had not been presented previously. Each scene was presented for 3.5s, as in the study phase, and was subsequently replaced by a recognition judgment screen. Only the old scenes were used in primary analyses; the new scenes served as recognition lures.

For the recognition judgment, participants indicated whether or not they recognized the scene from the study phase. They were given as much time as they needed to select their response. Response options fell on a 1-5 and recollect scale made up of “sure new,” “maybe new,” “don’t know,” “maybe old,” “sure old,” and “recollect old” (Ramey et al., 2019; Yonelinas, 2002). Participants were instructed and tested on how to use this scale prior to beginning the test phase.

Data Reduction and Analysis

Meaning maps. The meaning maps used were those created in Henderson and Hayes (2017), in which participants recruited via Amazon Mechanical Turk rated the meaningfulness of overlapping image patches of varying sizes. Specifically, they rated how informative or recognizable the visual information contained in each patch was. For each scene, the patch ratings were used to construct a map of the spatial distribution of meaning (Figure 1c, Figure 5c). For more details on how the meaning maps were generated, see Henderson and Hayes (2017). The resulting map for each scene was a 1024x768 matrix, with each cell corresponding to a pixel of the scene. The value in each cell represents the intensity of meaning at that point in the scene. The maps were gaussian smoothed to account for the fall-off in visual acuity from the fovea.

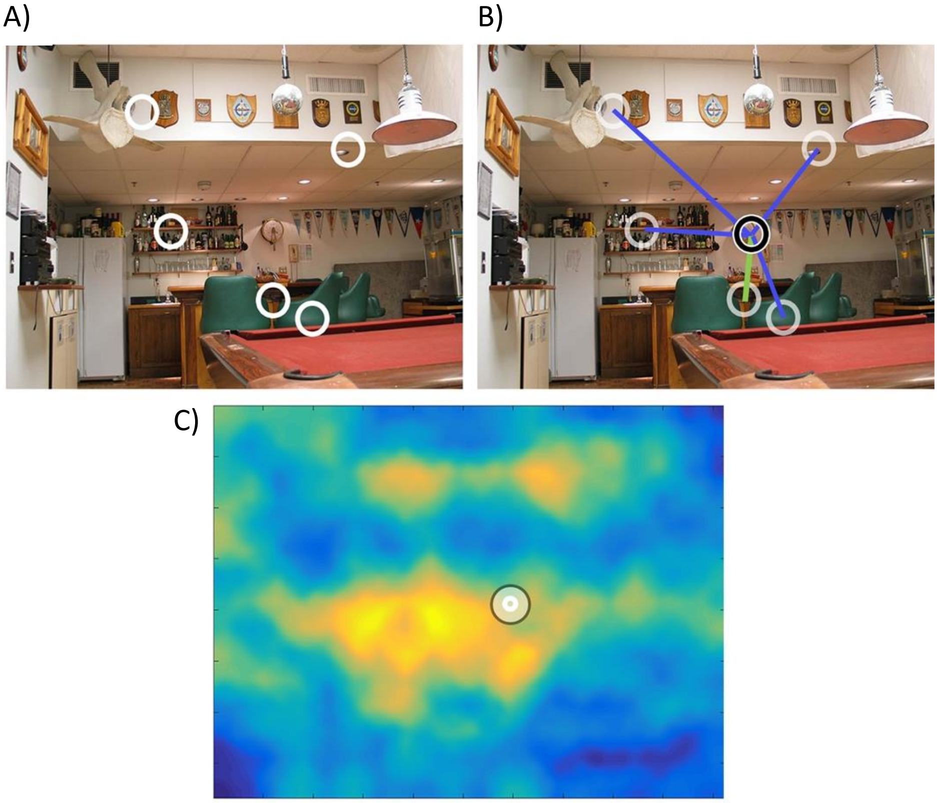


Figure 5. Resampling and meaning measures. A) Fixations (white rings) made while studying the scene. B) One of the fixations made while viewing the scene during test (black ring) along with the study fixations from a) (white rings). To calculate refixation distance for each test fixation, we computed the distance (lines) between the test fixation and every fixation made during study of that scene. The shortest resulting distance was assigned as the refixation distance score for that test fixation (green line). Thus, a lower refixation distance indicates that a test fixation was nearer to a region visited during study. C) The meaning map of the scene. The brighter, yellow regions denote areas of high meaning, whereas the dark blue regions denote areas of low meaning. The meaning score of each test fixation was calculated by taking the average of the density of meaning within a 1° radius (black ring) around the fixation coordinates. The size of the 1° radius black ring is drawn to scale; the fixations, however, are recorded as a single coordinate and are not drawn to scale.

Eye Movements. Fixations and saccades were segmented with EyeLink's standard algorithm using velocity and acceleration thresholds (30°/s and 9500°/s^2; SR Research, 2010b). Eye movement data were imported offline into Matlab using the EDFConverter tool. The first fixation was excluded from all analyses because its location was determined by the experiment-defined central fixation point.

Resampling. We computed resampling on a fixation-by-fixation basis in order to determine the extent to which participants visited regions during test that were the same as (or near) those visited during study (Figure 5a-b). To do this, we considered each fixation made during the test phase individually. For each test phase fixation on each scene by each subject, we computed the distance to each study phase fixation on that same scene by that same subject. The shortest resulting distance was reserved for analysis, and was termed *refixation distance*: the distance from a test fixation to the nearest region that had been viewed during study. Refixation distance thus measures the extent to which a test fixation was far from any previously visited region, such that a lower refixation distance reflects more resampling behavior.

Meaning scores. The extent to which participants attended to meaningful regions was calculated on a fixation-by-fixation basis, similar to refixation distance (Figure 5c). The meaning map for a given scene was used to compute the average amount of meaning contained in a 1-degree radius around each test fixation on that scene. This yielded a meaning score for each fixation.

Statistical models. Statistical analyses were conducted using linear mixed effects models, which allowed us to harness trial-by-trial (i.e., within-subjects) data while controlling for individual differences and stimulus effects (Nuthmann and Einhäuser 2015). In addition to random intercepts of subject and image for all analyses, we nested fixations within trials for fixation-by-fixation analyses. The dependent variable in each model of resampling was refixation distance. The models were estimated using the lmerTest package in R (Kuznetsova, Brockhoff, & Christensen, 2017), and were fit using maximum likelihood. The degrees of freedom and *t* values used were output by the linear mixed effects model for the variables of interest. The degrees of freedom were computed using the Satterthwaite approximation, and were rounded to the nearest integer in the manuscript. Effect sizes were calculated as a standardized regression coefficient (β).

References

Castelhano, M. S., Mack, M. L., & Henderson, J. M. (2009). Viewing task influences eye movement control during active scene perception. Journal of Vision, 9(3), 6. http://dx.doi.org/10.1167/9.3.6Damiano C, Walther DB. 2015. Content, not context, facilitates memory for real-world scenes. *Vis cogn* **23**: 852–855. https://www.tandfonline.com/action/journalInformation?journalCode=pvis20 (Accessed July 17, 2019).

Damiano C, Walther DB. 2019. Distinct roles of eye movements during memory encoding and retrieval. *Cognition* **184**: 119–129. https://www.sciencedirect.com/science/article/pii/S0010027718303305 (Accessed January 7, 2019).

Foulsham T, Kingstone A. 2013. Fixation-dependent memory for natural scenes: An experimental test of scanpath theory. *J Exp Psychol Gen* **142**: 41–56. https://search.proquest.com/psycinfo/docview/1011297796/fulltextPDF/6252F0D9737B484APQ/1?accountid=14505 (Accessed September 28, 2017).

Godden DR, Baddeley AD. 1975. Context-dependent Memory in Two Natural Environments: on Land and Underwater. *Br J Psychol* **66**: 325–331. https://psycnet.apa.org/record/1978-22375-001 (Accessed October 19, 2019).

Harel J, Koch C, Perona P. 2006. Graph-Based Visual Saliency. *Adv Neural Inf Process Syst*.

Henderson JM. 2003. Human gaze control during real-world scene perception. *Trends Cogn Sci* **7**: 498–504. http://jhenderson.org/vclab/PDF_Pubs/Henderson_TICS_2003.pdf (Accessed January 23, 2018).

Henderson JM, Hayes TR. 2017. Meaning-based guidance of attention in scenes as revealed by meaning maps. *Nat Hum Behav* **1**: 743–747. http://www.nature.com/articles/s41562-017-0208-0.

Henderson JM, Hayes TR. 2018. Meaning guides attention in real-world scene images: Evidence from eye movements and meaning maps. *J Vis* **18**: 10. http://jov.arvojournals.org/article.aspx?doi=10.1167/18.6.10 (Accessed July 30, 2019).

Holm L, Mantyla T. 2007. Memory for scenes: Refixations reflect retrieval. *Mem Cogn* **35**: 1664–1674.

Humphrey K, Underwood G. 2010. The potency of people in pictures: Evidence from sequences of eye fixations. *J Vis* **10**: 19–19. http://jov.arvojournals.org/Article.aspx?doi=10.1167/10.10.19 (Accessed July 16, 2018).

Itti L, Koch C. 2001. Computational modelling of visual attention. *Nat Rev Neurosci* **2**: 194–203. http://www.nature.com/articles/35058500 (Accessed September 23, 2018).

Kardan, O., Berman, M. G., Yourganov, G., Schmidt, J., & Henderson, J. M. (2015). Classifying mental states from eye movements during scene viewing. Journal of Experimental Psychology: Human Perception and Performance, 41, 1502–1514. http://dx.doi.org/10.1037/a0039673

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest Package: Tests in linear mixed effects models. Journal of Statistical Software, 82, 1–19. http://dx.doi.org/10.18637/jss.v082.i13

Locher PJ, Nodine CF. 1974. The role of scanpaths in the recognition of random shapes. *Percept Psychophys* **15**: 308–314. http://www.springerlink.com/index/10.3758/BF03213949 (Accessed September 16, 2019).

Mancas M, Le Meur O. 2013. Memorability of natural scenes: The role of attention. In *2013*

IEEE International Conference on Image Processing, pp. 196–200, IEEE http://ieeexplore.ieee.org/document/6738041/ (Accessed July 30, 2019).

Mantyla T, Holm L. 2006. Gaze control and recollective experience in face recognition. *Vis cogn* **14**: 365–386. http://www.tandfonline.com/doi/pdf/10.1080/13506280500347992?needAccess=true (Accessed September 28, 2017).

Mills, M., Hollingworth, A., Van der Stigchel, S., Hoffman, L., & Dodd, M. D. (2011). Examining the influence of task set on eye movements and fixations. Journal of Vision, 11(8), 17. http://dx.doi.org/10.1167/11.8.17

Neider, M. B., & Zelinsky, G. J. (2006). Scene context guides eye movements during visual search. Vision Research, 46(5), 614–621. https://doi.org/10.1016/j.visres.2005.08.025

Noton D, Stark L. 1971. Scanpaths in eye movements during pattern perception. *Science* **171**: 308–11. http://www.ncbi.nlm.nih.gov/pubmed/5538847 (Accessed April 6, 2018).

Nuthmann A, Einhäuser W. 2015. A new approach to modeling the influence of image features on fixation selection in scenes. *Ann N Y Acad Sci* **1339**: 82–96.

Ramey MM, Yonelinas AP, Henderson JM. (2019). Conscious and unconscious memory differentially impact attention: Eye movements, visual search, and recognition processes. *Cognition* **185**: 71–82. https://psyarxiv.com/p8qre (Accessed January 12, 2019).

Ramey, M. M., Henderson, J. M., & Yonelinas, A. P. (2020). The spatial distribution of attention predicts familiarity strength during encoding and retrieval. *Journal of Experimental Psychology: General.*

Ryan, J. D., & Shen, K. (2020). The eyes are a window into memory. In Current Opinion in Behavioral Sciences (Vol. 32, pp. 1–6). Elsevier Ltd.

https://doi.org/10.1016/j.cobeha.2019.12.014

Saffran M, Schwartz F. 1994. Of cabbages and things: Semantic memory from a neuropsychological perspective--A tutorial review. *Atten Perform XV Conscious nonconscious Inf Process* **25**: 507–536. https://psycnet.apa.org/record/1994-98352-020 (Accessed November 19, 2019).

Smith SM. 1979. Remembering in and out of context. *J Exp Psychol Hum Learn Mem* **5**: 460–471.

SR Research (2010a). Experiment builder user’s manual. Mississauga, Ontario, Canada: SR Research Ltd.

SR Research (2010b). EyeLink 1000 user’s manual (version 1.5.2). Mississauga, Ontario, Canada: SR Research Ltd.

Tatler BW, Hayhoe MM, Land MF, Ballard DH. 2011. Eye guidance in natural vision: Reinterpreting salience. *J Vis* **11**: 5–5. http://jov.arvojournals.org/Article.aspx?doi=10.1167/11.5.5 (Accessed July 23, 2018).

Tulving E. 1986. Episodic and semantic memory: Where should we go from here? *Behav Brain Sci* **9**: 573–577.

Valuch C, Becker SI, Ansorge U. 2013. Priming of fixations during recognition of natural scenes. *J Vis* **13**: 1–22. http://jov.arvojournals.org/Article.aspx?doi=10.1167/13.3.3 (Accessed March 6, 2018).

Van der Stigchel S, Belopolsky A V., Peters JC, Wijnen JG, Meeter M, Theeuwes J. 2009. The limits of top-down control of visual attention. *Acta Psychol (Amst)* **132**: 201–212. https://www.sciencedirect.com/science/article/pii/S0001691809000900?via%3Dihub (Accessed August 7, 2018).

Võ ML-H, Wolfe JM. 2013. The interplay of episodic and semantic memory in guiding repeated search in scenes. *Cognition* **126**: 198–212.

Wynn JS, Bone MB, Dragan MC, Hoffman KL, Buchsbaum BR, Ryan JD. 2016. Selective scanpath repetition during memory-guided visual search. *Vis cogn* **24**: 15–37. http://www.tandfonline.com/action/journalInformation?journalCode=pvis20 (Accessed September 28, 2017).

Wynn JS, Olsen RK, Binns MA, Buchsbaum BR, Ryan JD. 2018. Fixation reinstatement supports visuospatial memory in older adults. *J Exp Psychol Hum Percept Perform* **44**: 1119–1127. http://doi.apa.org/getdoi.cfm?doi=10.1037/xhp0000522 (Accessed March 6, 2018).

Wynn, J. S., Ryan, J. D., & Moscovitch, M. (2020). Effects of prior knowledge on active vision and memory in younger and older adults. Journal of Experimental Psychology: General, 149(3), 518–529. https://doi.org/10.1037/xge0000657

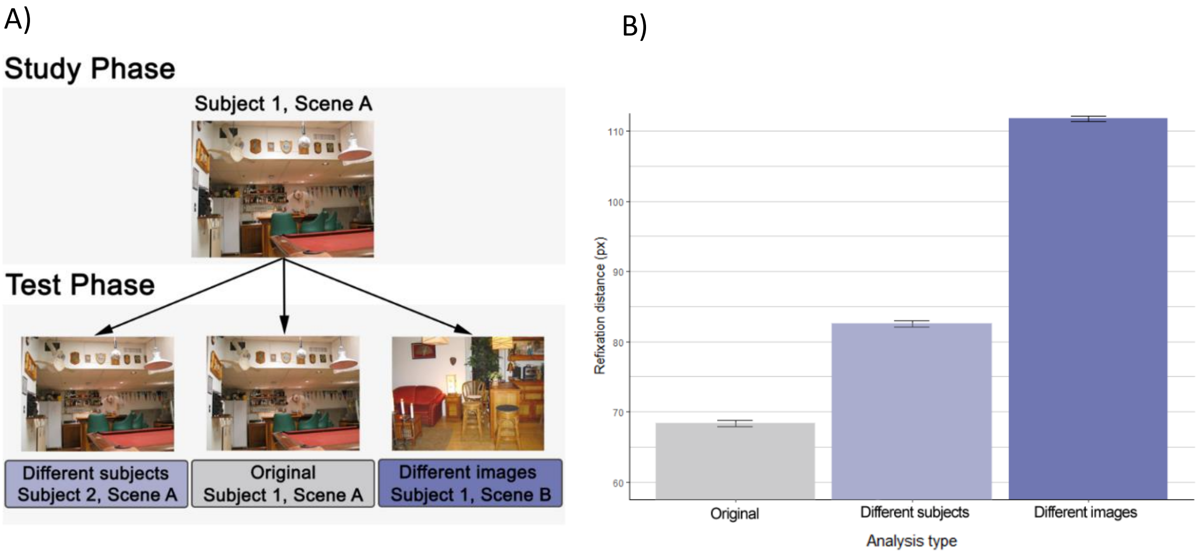
Wynn JS, Shen K, Ryan JD. 2019. Eye Movements Actively Reinstate Spatiotemporal Mnemonic Content. *Vision* **3**: 21. https://www.mdpi.com/2411-5150/3/2/21 (Accessed September 3, 2019).

Yonelinas AP. 2002. The nature of recollection and familiarity: A review of 30 years of research. *J Mem Lang* **46**: 441–517.

Appendix

Resampling Baselines

Our primary analyses were focused on how subjective recognition memory and the semantic meaning of fixated scene regions were related to resampling behavior. However, to ensure that the overall observed resampling behavior was above and beyond that predicted by individual differences in general viewing behavior or general scene properties, we examined resampling for random pairings of scene and subject (Supplemental Figure 1a). We did this by computing the refixation values across different subjects and across different items. That is, for the cross-subject analysis, we paired each subject's study phase data on a given scene with another subject's test phase data from the same scene, and calculated refixation distance for that pair. Any "resampling" in this case could not be driven by an individual subject's unique memory for the scene, but rather is likely driven by shared scene properties such as meaning, salience, and the overall configuration of the scene (e.g., a scene may have information concentrated on its right half with a relatively empty left half, thus drawing attention to the right on average). Similarly, for the cross-image analysis, we paired each subject's study phase data from a given scene with the test phase data from a different scene, but from that same subject. Any "resampling" in this case is likely driven by global similarities in viewing behavior, such as center bias (i.e., a subject's tendency to attend to the center of the screen). This process provided two baselines to which we could compare the observed resampling behavior. The results indicate that the correctly paired (i.e., actual) resampling behavior was substantially higher than either the cross-subject, $t(5312)= 14.5$, $p<.0001$, $d= 0.40$, or the cross-image pairings, $t(4902)= 40.6$, $p<.0001$, $d= 1.16$ (Supplemental Figure 1b). This suggests that both subject-level factors (e.g., episodic memory) and image-level factors (e.g., meaningful scene regions) did lead to increased resampling behavior. Furthermore, the larger reduction in resampling behavior from the cross-image analysis (i.e., higher refixation distance; Supplemental Figure 1b) suggests that image properties such as salience, meaning, and overall distribution of image content may be responsible for the majority of resampling behavior.

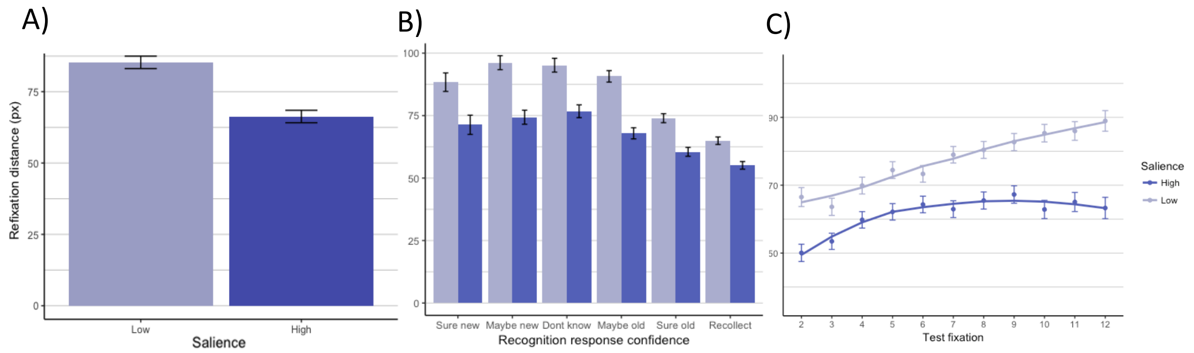


Supplemental Figure 1. Refixation distance for random pairings of subject and image. Note that a lower refixation distance indicates more resampling behavior. A) The method of calculating refixation distance using random pairings. Study phase data from a given subject viewing a given scene was randomly paired with the test data from either a different subject viewing the same scene, or the same subject viewing a different scene. The refixation distance values computed using these pairings were compared to the values computed using the correct pairing of same subject with same scene (i.e., original pairing) B) The refixation distance results. The "original" bar represents the actual pairings of trials (i.e., no random pairings); the "different subjects" bar is the cross-subject analysis such that different subjects viewing the same scene were paired; and the "different images" bar is the cross-image analysis such that different scenes were paired within the same subject.

Salience Analyses

One prominent theory of attentional guidance proposes that where we choose to attend, in general, is driven by bottom-up visual salience such as edges and high contrast (Itti and Koch 2001). Salience is conceptually related to meaning in that they are converging metrics of the informativeness of scene regions, but salience lacks the ability to serve as an index of guidance by semantic memory. Furthermore, recent studies have indicated that meaning and salience are highly correlated, and meaning has been shown to account for much, if not all, of the ability of salience to predict attention (Henderson and Hayes 2017, 2018)—and meaning is a better overall predictor of attention. Nevertheless, to increase the interface of the present findings with prior studies, and for those investigators interested in salience, we ran salience through the same set of analyses as we did meaning.

Computing salience scores. Saliency maps were computed using the Graph Based Visual Saliency (i.e., GBVS) model for combining low-level feature maps (e.g., orientation, color) into a map of salient regions in an image (Harel, Koch, & Perona, 2006). It should be noted that the standard practice for computing saliency maps is to apply a center bias (i.e., down-weighting of values in the periphery) in order to increase predictive accuracy of saliency models of attention, and this process is included in the GBVS implementation. However, because we were interested in salience in terms of low-level image features but not in predicting attention *per se*, we included a covariate to eliminate center bias in all analyses of salience. To do this, the distance from each test fixation to the center was computed, and was controlled for in salience analyses. Just as with meaning scores, salience scores were computed on a fixation-wise basis, by determining the average amount of salience contained in the saliency map of the region immediately surrounding each fixation.



Supplemental Figure 2. Salience analyses. Least-squares means derived from the linear mixed effects models used in analysis are plotted, and the error bars represent the standard error of these estimated means from the model. Center bias is controlled for in the plots. A) Refixation distance by attention to salience (i.e., salience score) B) Interaction between salience and memory response. C) Refixation distance over fixations by the salience score of each fixation.

Results. To determine whether salience related to resampling behavior, we regressed the refixation distance of each fixation on the salience score of each fixation in the test phase. Salience scores significantly predicted refixation distance, β= -.29, *p*<.0001, indicating that the tendency to resample a region was associated with the salience of that region (Supplemental Figure 2a).

Interaction between salience and episodic memory. To determine whether salience modulated the relationship between episodic memory and refixation distance, we regressed refixation distance on the interaction between salience score and memory strength. A nonsignificant interaction trend was observed between salience and memory strength, β= .02, *p*=.083 (Supplemental Figure 2b), which was in the same direction as that observed using meaning.

***Temporal analysis*.** Similar to meaning, the relationship between salience and refixation distance changed significantly over fixations, β = -.02, *p*<.0001, such that salience had a larger effect on resampling later in the trial (Supplemental Figure 2c). In contrast to the memory strength results, there was no significant quadratic effect, *p*>.25.

The results of the salience analyses provide converging evidence that the informativeness of scene regions, whether due to bottom-up visual information or top-down semantic knowledge, may drive resampling behavior.

Combined Models

To determine whether all three factors (i.e., memory strength, meaning, and salience) still contributed to resampling behavior when considered together, we included all three variables in a model predicting refixation distance (along with the covariate of distance from the center; Eq. A1). Note that this model and all of the models in this section had two random effects: a random intercept of image, and a random intercept of trial nested within subject, as in the main text. This controls for subject and image-level effects, and ensures that fixations are nested within trials (given that fixations can be thought of as repeated measures within a trial).

In this combined model, memory strength continued to be a significant predictor of refixation distance, $\beta = -.12$, $p < .0001$, as did salience, $\beta = -.10$, $p < .0001$, and meaning, $\beta = -.16$, $p < .0001$. This indicates that all three variables predicted unique variance in resampling behavior that was not due to center bias.

Eq. A1

Refixation distance ~ memory + meaning + salience + distance from center

For the sake of completeness, we also examined this combined model with the interaction term of memory by meaning (Eq. A2) to determine if it improved the model fit compared to the model examined above (Eq. A1). (Note that the main effects of memory and meaning were still included in Eq. A2.) An ANOVA between these two models confirmed that adding the interaction term of memory strength by meaning accounted for variance in refixation distance above and beyond the model in Eq. A1, *p*<.0001.

Eq. A2

Refixation distance ~ **memory x meaning** + salience + distance from center

To further confirm that each predictor accounted for unique variance in refixation distance, we conducted a model comparison for each predictor to determine whether it significantly improved a model containing the other two predictors (and a covariate of center bias). For example, for salience, we examined whether the addition of salience as a predictor significantly improved a model predicting refixation distance from meaning, memory strength, and distance from the center (Eq. A3). Specifically, in the Equation A3 example below, model A) was compared to model B) to determine whether salience predicted variance in refixation distance above and beyond that predicted by the other three variables. This process was repeated for each variable (Eq. A4, Eq. A5). Conducting ANOVAs between the linear mixed effects models confirmed that memory strength, salience, and meaning each uniquely improved the ability to predict refixation distance, *p*s<.0001. This indicates that these three factors each accounted for unique variance in resampling behavior, and this was not due to center bias.

Eq. A3

A) Refixation distance ~ memory + meaning + distance from center

B) Refixation distance ~ **salience** + memory + meaning + distance from center

Eq. A4

A) Refixation distance ~ salience + meaning + distance from center

B) Refixation distance ~ **memory** + salience + meaning + distance from center

Eq. A5

A) Refixation distance ~ salience + memory + distance from center

B) Refixation distance ~ **meaning** + salience + memory + distance from center

Robustness checks

Controlling for number of fixations. Resampling behavior, and therefore refixation distance, is intuitively related to the number of fixations made at study ($\beta = -.21$, $p<.0001$): More fixations at study increases the chance that any given test fixation will be near a study fixation. Therefore, we re-ran the refixation distance analyses with a covariate of the number of fixations made during study for a given scene and subject, to ensure that the effects obtained were not driven by this relationship. All significant primary resampling effects remained significant, *p*s<.0001, and all temporal effects remained significant, *p*s<.01, with a covariate of the number of fixations made during study (including salience effects).

Controlling for center bias in meaning analyses. A covariate of distance from the center was included in all salience analyses to account for the center bias that is introduced in the process of generating salience maps. Although meaning maps do not have this confound, it is possible that meaningful information may naturally be biased towards the center of scenes. To ensure this did not influence the meaning results obtained, we re-ran all meaning analyses with a covariate of distance from the center. All results held, *p*s<.001.

Including recollect responses. Recollect responses were included on the scale for consistency with our prior studies (Ramey et al. 2019, 2020), but were not included in the above analyses given our *a priori* focus on familiarity strength. However, for the sake of completeness, we re-ran the analyses with recollect responses included to ensure this did not alter the results. All effects of memory response on resampling held, as did the interaction with meaning, *p*s<.001. The only difference that emerged was that salience was no longer marginal in its interaction with memory to predict resampling, but was significant, β= .03, *p*<.0001 (Supplemental Figure 2b).

Matching the meaning and salience maps. The salience and meaning maps differed in terms of the overall density of meaning and salience in the scenes. Although analyses involved comparing how the variation *within* the meaning and salience scores of scenes predicted refixation distance, it is possible that the systematic differences in density could bias results in the model containing both salience and meaning. To ensure that this was not the case, we re-ran all analyses using meaning and salience maps that were matched in density to the fixation map for a given trial. All results held, *p*s<.001.

Collapsing across subjects. We used linear mixed effects models in the manuscript to harness trial-by-trial and fixation-by-fixation data, but we re-ran the memory analysis by subject to ensure that the observed memory effects were not biased by different numbers of trials in different memory response bins. Specifically, memory strength and refixation distance were averaged by subject and submitted to a standard linear regression. We found that subjects with higher memory confidence for old scenes demonstrated more resampling behavior, *p*=.013, and that subjects with a higher *d*’ demonstrated more resampling behavior, *p*=.015.

Additional Analyses

Relationship between meaning, memory, and salience. In order to further confirm that these forms of guidance may be considered to be independently involved in resampling (i.e., that they were not mediated by each other), we assessed the correlations between each variable during the test phase. Note that these correlations (Eq. A6) were distinct from models in the main text that assessed resampling behavior (Eq. A7); we clarify this distinction in the equations below. Memory response was not related to the meaning, $\beta=.005$, $p=.29$, nor salience, $\beta=.006$, $p=.45$, scores of fixations. (Note that the nesting of fixations within trials accounted for the comparison of a trial-level measure with a fixation-level measure.) As expected based on prior literature finding correlations between meaning and salience of regions (Henderson and Hayes 2017, 2018), meaning and salience scores were significantly associated, $\beta=.28$, $p<.0001$. Thus, it appears that memory strength is independent from meaning and salience at test, and they may have the potential to comprise independent forms of guidance.

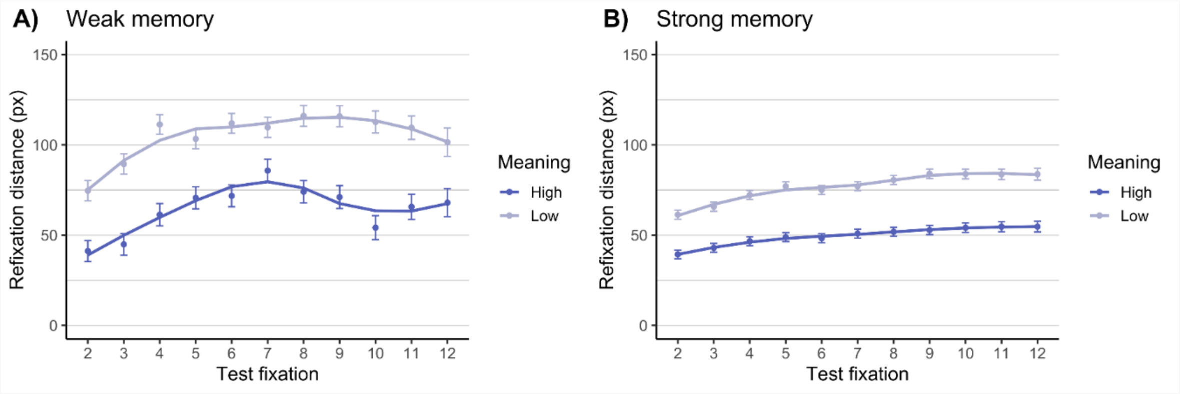
Eq. A6: Relationship between meaning and memory

Meaning ~ memory

Eq. A7: Interaction between meaning and memory in predicting resampling

Refixation distance ~ meaning x memory

Interaction between semantic meaning and episodic memory over time. To determine whether the negative interaction between meaning and memory observed in the main text (Figure 2d) varied over time, we examined the three-way interaction between meaning, memory, and ordinal fixation number (Supplemental Figure 3). There was no significant three-way interaction when fixation number was examined linearly or with a quadratic function, *p*s>.39. This indicates that the interaction between meaning and memory did not vary systematically over the course of a trial.

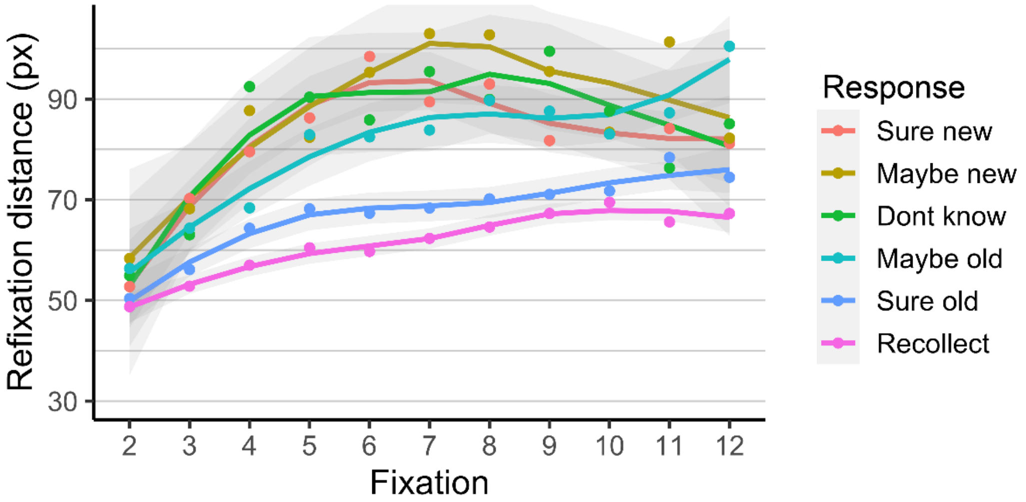


Supplemental Figure 3. The interaction between meaning and memory in predicting refixation over the course of test fixations. Least-squares means derived from the linear mixed effects models used in analysis are plotted, and the error bars represent the standard error of these estimated means from the model. A) Weak memory responses included “Sure new” and “Maybe new” responses, B) Strong memory responses included “Sure old” and “Recollect” responses.

Study phase analysis. We hypothesized that salience and meaning would lead to resampling behavior because they may guide attention during study, and do the same during test—thus leading participants to view similar regions across viewings. However, the analyses thus far have only examined how viewing salient and meaningful regions *during test* relates to resampling. While it logically follows from the results that attention to such regions during study would also relate to resampling, it is not clear whether increased attention to salient and meaningful regions during study might strengthen the relationship between resampling and viewing such regions during test. To directly examine this, we correlated the fixation maps (i.e., smoothed heatmaps of fixation locations) made during study of each image by each subject with the salience and meaning maps to produce a trial-level version of the salience and meaning scores used above. We then determined whether the salience and meaning correlations of trials during study interacted with the salience and meaning scores of fixations during test to predict resampling. Note that the study phase correlation for a subject and image applied to all test phase fixations for that same image viewed by that subject, as the goal of the measure was to determine whether *overall* viewing of salient or meaningful regions during study of an image affected how salience or meaning scores for fixations on that image at test were related to resampling. To test this idea, we used a linear mixed effects model that regressed refixation distance on the interaction between the test salience scores of fixations and the trial-wide salience correlations at study—then re-ran the model with test meaning scores and study meaning correlations.

We predicted that increased viewing of salient or meaningful regions in an image during study would synergistically interact with attention to salience or meaning while viewing that image during test to predict resampling. For example, increased attention to meaning during test should only predict resampling if people also viewed meaningful regions during study. We found that this was the case for both salience and meaning. Specifically, viewing salient regions during study interacted with viewing salient regions during test to predict decreased refixation distance, $\beta = -.14$, $p<.0001$, indicating that the relation between test salience scores and resampling was stronger when the study salience correlation was higher. The same effect was found with respect to meaning, $\beta = -.15$, $p<.0001$, such that the relation between test meaning scores and resampling was stronger when the study meaning correlation was higher. These results demonstrate that viewing more salient and meaningful regions at study increases the extent to which viewing them at test is related to resampling, and thus indicate that salience and meaning lead to resampling behavior by guiding the eyes to the same regions at both study and test.

Additional Plot



Supplemental Figure 4. Refixation distance over test fixations by memory response. The x-axis represents the ordinal fixation number in a trial. Least-squares means derived from the linear mixed effects models used in analysis are plotted, and the error shading represents the standard error of these estimated means from the model. The lines were generated using a loess smoothing (i.e., locally weighted smoothing) function.

Chapter 4

Schema knowledge and episodic memory trade off to influence spatial memory

Michelle M. Ramey[a,b,c*], John M. Henderson[a,c] Andrew P. Yonelinas[a,b]

[a] Department of Psychology, University of California, Davis, CA, USA.

[b] Center for Neuroscience, University of California, Davis, CA, USA.

[c] Center for Mind and Brain, University of California, Davis, CA, USA.

Author Note:

Funding: This work was supported by the National Eye Institute of the National Institutes of Health under Award Numbers R01EY025999 and R01EY027792. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Abstract

Schema knowledge can dramatically affect how we encode and consolidate new memories, but how schemas impact memory retrieval is largely unknown. Although schemas and memory are thought to operate in conjunction, recent neurocognitive models suggest that they may instead compete. To test this possibility, we had participants search for target objects in semantically expected (i.e., congruent) locations within scenes or in unusual (i.e., incongruent) locations. In a subsequent test, participants indicated where in each scene the target had been located previously, then provided confidence-based recognition memory judgments that indexed recollection, familiarity strength, and unconscious memory for the scenes. In two studies (N=210), spatial memory was more accurate for schema-congruent than incongruent locations, but this effect weakened with more strongly recognized scenes, and was eliminated for recollected scenes. The results support competition models and indicate that schema knowledge contributes to spatial memory primarily when recognition memory fails to provide precise information.

Keywords: Recognition, semantic knowledge, schemas, recognition memory, spatial memory

Schema knowledge and episodic memory trade off to influence spatial memory

Each day we make predictions about where objects will be in the world around us. These predictions are often based on memory for past experiences: We can find an object, such as our keys, by remembering where the object was the last time we saw it. Alternatively, we can draw on our knowledge about where objects are *typically* located, using semantic knowledge in the form of schemas (Bartlett & Burt, 1932; Ghosh & Gilboa, 2014; Henderson, 2003; Rumelhart, 2017). Although schemas can often facilitate performance (Alba & Hasher, 1983; Anderson, 1981; Bartlett & Burt, 1932; Greve et al., 2019; Rumelhart, 2017), they can also lead to memory impairments and distortions (Johnson, 1997; Lew & Howe, 2017; Roediger & McDermott, 1995; Sweegers et al., 2015), and their inappropriate application has been linked to catastrophic aviation and driving errors (Dismukes, 2008; Hole, 2014; Plant & Stanton, 2012) as well as prejudice and depression (Cox et al., 2012; Hawke & Provencher, 2011). But how are schemas and episodic memory combined to guide our behavior? Despite the relevance of this question for daily life and the rapidly growing scientific interest in schemas (e.g., Gilboa & Marlatte, 2017), it is not yet understood how schemas and memory interact to influence behavior (Gilboa & Marlatte, 2017; van Buuren et al., 2014; van Kesteren et al., 2012).

Prior work on schemas and memory has indicated that the hippocampus is critical in supporting memory for individual episodes, whereas the neocortex and the medial prefrontal cortex are involved in supporting schema knowledge (Gilboa & Marlatte, 2017; McClelland et al., 1995; Robin & Moscovitch, 2017; van Kesteren et al., 2010). It has been proposed that these systems function in a complementary manner (McClelland, 2013; McKenzie et al., 2013, 2014; Preston & Eichenbaum, 2013) with schema knowledge and memory synergistically contributing to a desired behavioral outcome. Based on this proposal, behavioral performance is expected to be better when both forms of information are available than either one alone.

In contrast, however, recent work has suggested that schemas and episodic memory may instead operate in a competitive or inhibitory manner (Gilboa & Marlatte, 2017; Sommer, 2017; van Kesteren et al., 2012; Wagner et al., 2015). For example, whereas novel learning is generally related to hippocampal activity, schema-related learning has been associated with a shift to medial prefrontal involvement (Sommer, 2017; Wagner et al., 2015). When schema knowledge becomes more relevant, cortical representations are engaged, and the hippocampus appears to be suppressed (Bein et al., 2014; Gilboa & Marlatte, 2017; van Kesteren et al., 2012, 2013). If these neural findings are reflected in behavior, schema knowledge and episodic memory may compete to influence performance. Based on this competition hypothesis, therefore, one might expect that when memory is stronger, schema knowledge may contribute less to behavioral performance, and vice versa.

Despite the support for each of these possibilities, they have not yet been directly examined. There are, however, some studies indirectly examining how schema knowledge and memory interact. For example, Võ & Wolfe (2013) found that memory-driven improvements in search speed across repeated presentations were more dramatic when semantic information was unavailable. Moreover, in an eyetracking study, we found that attention in scenes was less driven by semantic meaning when participants’ familiarity with the scenes was stronger (Ramey, Yonelinas, et al., 2020). Although the latter study did not examine schemas, these results suggest that semantic knowledge may interact with memory such that its influence decreases when memory strength increases.

The present study aimed to directly address the question of how schema knowledge and episodic memory strength interact to influence behavioral performance. To do this, we examined the impact of spatial schema knowledge and recognition memory strength for scenes on the ability to determine where a target object was located in a previously viewed scene. First, participants searched for target objects in scenes during an initial study phase. Half the scenes contained the target object in a schema-congruent location (e.g., toothbrush next to sink), and half contained the target object in a schema-incongruent location (e.g., toothbrush next to bathtub; see Figure 1). During a subsequent test phase, participants were shown a mixture of new and studied scenes that did not contain the target object and were asked to indicate the precise location that had contained the target object during the study phase. To examine the joint effects of schema congruence and memory strength for the scene on spatial accuracy, participants made a recognition memory response on a confidence-based scale that allowed us to examine familiarity strength, recollection, and unconscious memory for the scenes (Ramey et al., 2019). If memory and schema knowledge compete, then schema congruency effects should decrease in magnitude as familiarity confidence increases. In addition to providing familiarity confidence, participants indicated if the scenes were confidently recognized *and* were accompanied by an ability to remember specific details about the study event (i.e., recollection; e.g., Tulving, 1982, 1985). Recollect responses have been interpreted either as reflecting a distinct hippocampus-based recollection process, or as reflecting extremely high levels of episodic memory strength (e.g., see Eichenbaum et al., 2007; Wixted, 2007; Yonelinas, 2002). In either case, if schema knowledge competes with episodic memory, then the schema congruency effects should be smallest for the recollect responses. Lastly, the confidence scale also allowed us to assess whether unconscious memory modulated the effects of schema congruence on spatial accuracy.

Unconscious memory was indexed by comparing performance on new scenes to performance on scenes that had been studied but that participants were highly confident were new (i.e., high-confidence misses; Ramey, Henderson, & Yonelinas, 2020; Ramey et al., 2019). We also replicated the initial experiment in a separate sample to assess the robustness of the results.

Method

Participants

One-hundred and fifty undergraduate students completed the experiment for course credit. A sample size of 120 participants was needed to provide 95% power to detect the most relevant previously observed effect, which we obtained in a prior study (i.e., the interaction between familiarity strength and semantic meaning in influencing attention; Ramey, Yonelinas, et al., 2020). We selected a final sample size of 150 participants to ensure that we would have adequate power even after potential exclusions. Participants were removed from analysis for technical issues or for failing to properly complete the task; specifically, due to clicking on the objects during the study phase less than 90% of the time (6 participants), quickly pressing a key to get through the recognition response phase (4 participants), or having atypical mouse coordinates that did not conform to the typical browser output (e.g., from using a tablet rather than computer; 7 participants). A total of 17 participants' data was removed, such that the sample used in analysis consisted of 133 participants.

The replication sample consisted of 60 undergraduate students who did not participate in the first experiment. The sample size was selected to provide 80% power to detect the familiarity strength effect in the first experiment. One participant was excluded for having atypical mouse coordinates, such that the final sample had 59 participants.

Materials

Apparatus. The study was conducted online using Javascript via JSpsych, which allows for accurate, high-speed presentation timing and response recording (de Leeuw, 2015). Participants were instructed to use a computer with a browser size of at least 800x600px. The experiment would not begin if a participant's browser size was less than 800x600px but allowed them to continue once they expanded it sufficiently; this requirement precluded use of a phone. Participants were able to see their cursor throughout the experiment.

Stimuli. Stimuli were 80 photographs of real-world scenes. All scenes were presented in color at a resolution of 800x600 pixels. Of these 80 scenes, 60 were presented at study and test (i.e., old scenes), and 20 were presented only at test (i.e., new scenes). We included more old scenes than new scenes to ensure that an adequate number of old scenes were recognized at each level of confidence for analysis. Stimulus presentation was counterbalanced, such that the scenes appeared in different conditions (i.e., presented at both study and test, or used as a new lure during test; see procedure) for different participants to mitigate stimulus effects.

Five scene categories were used, and a single type of target object was used for each category. The categories and targets consisted of kitchens (target: frying pan), dining rooms (target: wine glass), bedrooms (target: alarm clock), living rooms (target: coffee mug), and bathrooms (target: toothbrush cup). Eight different object exemplars were used per category, such that the visual features of the target object varied across different scenes within a category. In each scene, only one exemplar of the target object was present, and this was kept consistent across presentations. For example, in each living room scene, there was only one coffee mug present. Importantly, for a given scene viewed by a given participant, the target was always visually identical and in the same location across repeated viewings.

Two versions of each scene were created using Adobe Photoshop (Figure 1A-D): one with the target object in a schema-consistent location (i.e., *congruent* scene), and one with the target in an unexpected location (i.e., *incongruent* scene). The congruent location was consistent across all scenes in a category, such that targets were placed relative to larger objects with which the target objects co-occur with high probability in daily life (Boettcher et al., 2018; for review of scene grammar see Võ et al., 2019). Specifically, in bathroom scenes, the toothbrush cups were located next to sinks; in dining room scenes, the wine glasses were located on tables (within arm's reach of a chair); in kitchen scenes, the pans were on stove burners; in bedroom scenes, the alarm clocks were on night stands; and in living rooms, the coffee mugs were on coffee tables. In incongruent scenes, on the other hand, the objects were arbitrarily placed in unexpected but physically plausible locations (i.e., on floors, shelves, chairs, etc. rather than floating in space).

Scene congruence was manipulated within-subjects such that each participant was presented with half incongruent scenes and half congruent scenes. The congruent and incongruent versions of the scenes were also counterbalanced such that half of the participants saw the congruent version of a given scene, whereas the other half saw the incongruent version of that same scene. Importantly, a given scene was always congruent or incongruent within a given counterbalance, such that the target was always in the same place in a scene viewed across multiple repetitions by a given participant.

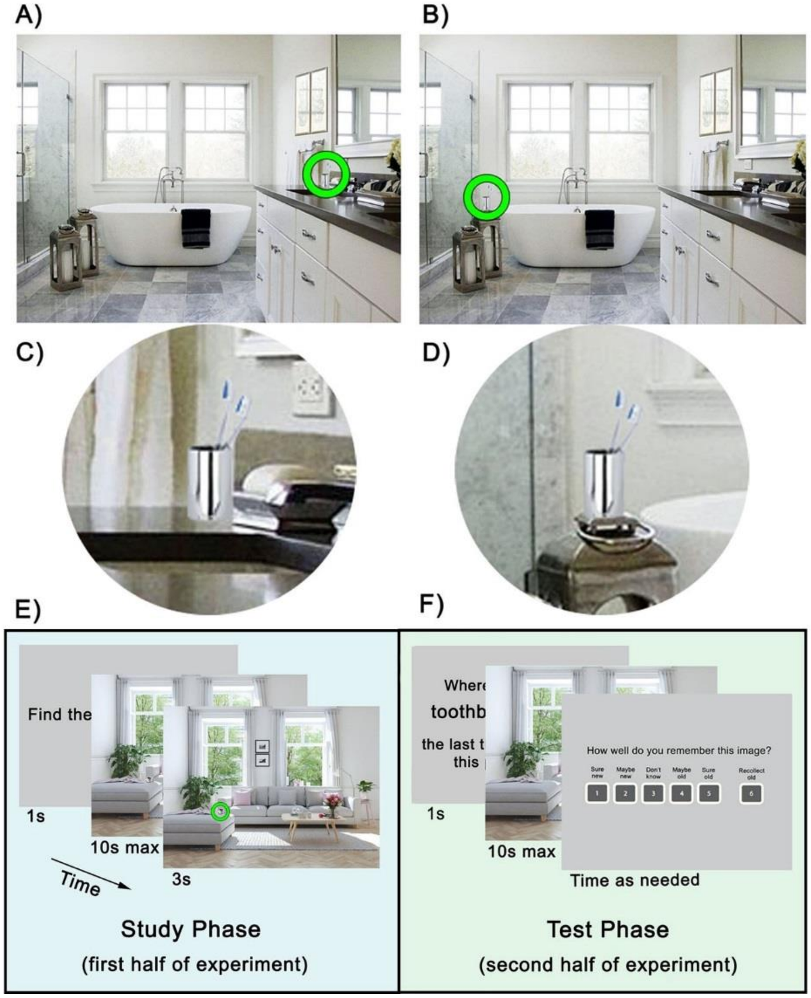


Figure 1. Sample stimuli and procedure. A) The congruent version of a sample scene, with the target object (toothbrush cup) next to the sink. The green ring appeared around the target after participants clicked on the scene in the study phase. B) The incongruent version of the scene. C) Closeup of the target object in the congruent scene (for visualization only; this was not part of the experiment). D) Closeup of the target object in the incongruent scene. E) The trial sequence in the study phase, which consisted of 60 scenes presented two times each (120 trials). In each trial, a target probe appeared (e.g., “Find the toothbrush cup”), followed by the scene with target object. Participants were required to click on the target object within 10s. After clicking or after 10s, whichever occurred first, a green ring appeared around the target for 3s. F) The trial sequence in the test phase, which consisted of 80 scenes (80 trials). A target probe appeared, followed by the scene without the target object, and participants were given 10s to click on the scene location that they thought had contained the target when the scene was presented in the study phase. After 10s or clicking, whichever occurred first, participants gave a confidence-based recognition memory response for the scene.

Procedure

The experiment lasted approximately 45 minutes and consisted of a study phase followed by a test phase (Figure 1E-F). There was a 2-minute break between the study and test phases.

Before each phase, participants were given instructions as well as three practice trials to familiarize them with the procedure. Participants were given a break midway through each phase. All procedures were approved by the University of California, Davis Institutional Review Board.

Study phase. Participants were told that they would be searching for and clicking on target objects and were asked to try to remember the scene and object locations for a later memory test. During the study phase, participants were presented with 60 unique scenes that were each repeated twice, for a total of 120 trials. The repetitions were randomly intermixed throughout the study phase, with the requirement that the same scene did not appear twice in a row. In each trial, participants were first given a 1s probe alerting them to the target object they would need to search for. For example, for dining room scenes, the probe was “Find the wine glass.” After the probe, the scene appeared, and participants had 10s to click on the target object in the scene. After clicking on the scene, or after 10s had elapsed, a green ring appeared around the target object and remained for 3s to allow participants to study the scene (Figure 1A-B).

Test phase. In the test phase, participants were asked to recall where the target object had been located in each scene when they had seen it during the study phase, and to provide a confidence-based recognition memory judgment for each scene. Participants were told that even if they thought that a scene was new (i.e., not presented in the study phase), they should make their best guess for where the target object might have been if it had in fact been in the study phase—that is, if their memory had failed and it actually was an old scene. The test phase included 80 scenes, 60 of which were presented in the study phase and 20 of which were new lures. Each scene was presented once for a total of 80 test trials. Each trial began with a target probe followed by the presentation of a scene without its target object, and participants were given 10s to click on the location in the scene where they remembered having seen the target object in the study phase. After clicking, or after 10s elapsed, a recognition memory response scale appeared and participants were given time as needed to respond.

Memory strength was measured by asking participants to rate memory confidence for each scene on a 6-point scale during the recognition judgment (Yonelinas, 2002). Participants were told that if they could consciously recollect some qualitative aspect of the initial learning event, such as what they thought about when the scene was encountered earlier, they should respond “Recollect old (6);” otherwise, they rated their memory confidence by responding “I’m sure it’s old (5),” “Maybe it’s old (4),” “I don’t know (3),” “Maybe it’s new (2),” or “I’m sure it’s new (1).” Importantly, participants were instructed that a “sure old” response was equal in confidence to a “recollect old” response, such that the only difference between them was that at least one specific detail of the learning event was remembered in recollected scenes. Participants were instructed and tested on how to use this scale prior to beginning the test phase.

Data Reduction and Analysis

The primary outcome of interest was *target distance:* the Euclidean distance between the location clicked by participants during the test phase and the actual location of the target object when the scene was presented in the study phase. This was measured in pixels between the mouse position during the click, recorded in terms of coordinates on the 800x600px scene, and the center of the target object. Reaction time during the study phase search task was also examined.

The effects of episodic memory were examined by comparing the target distance values between scenes given different recognition responses. Specifically, memory was compared across familiarity-based responses (1-5) to assess familiarity strength, and between recollected and “sure old” responses (6 versus 5) to assess recollection. In order to examine unconscious memory, performance needs to be compared to a memoryless baseline. To do this, we also computed target distance in scenes that were new. Participants’ clicks on new scenes represented their best guess for where the target object might have been located based on their schema knowledge and their knowledge of the experiment. From a participant’s point of view, these new scenes were equivalent to old scenes that they had forgotten seeing (i.e., high-confidence misses), but without any potential contribution of unconscious memory. To measure target distance in new scenes, we calculated the distance between the clicked location in the test phase and the target location when it was shown in the study phase for participants in the other counterbalance. In addition to providing a means by which to assess unconscious memory, this procedure allowed us to observe the effects of schema knowledge in the absence of episodic memory, and to determine whether episodic memory overall led to improvements in performance beyond those afforded by schema knowledge.

All statistical analyses besides those using new scenes were conducted using linear mixed effects models with random intercepts of subject and image, which allowed us to harness trial-by-trial (i.e., within-subjects) data while controlling for individual differences and stimulus effects. The models were estimated using the lmerTest package in R (Kuznetsova et al., 2017), and were fit using maximum likelihood. The degrees of freedom and *t* values used were output by the linear mixed effects model for the variables of interest. The degrees of freedom were computed using the Satterthwaite approximation, and were rounded to the nearest integer in the manuscript. For analyses using new scenes, standard linear regression was used because the measure was calculated between subjects (i.e., comparing performance between subjects who saw a given scene as a new scene and those who saw it as an old scene). Effect sizes were calculated as a standardized regression coefficient (β) for continuous variables, and Cohen's *d* for categorical variables.

Results

The proportion of trials that were given each type of recognition response is presented in Table 1. First, to visualize spatial memory accuracy, we plotted target location memory heatmaps for each condition (see Figure 2A-B). Each heatmap illustrates the distribution of recalled locations for scenes in that category, normalized across scenes such that the center of the heatmap represents the location of the target object. Heatmaps for schema-congruent and incongruent scenes were plotted for each type of recognition response (i.e., studied scenes given a recognition response of "recollect," "sure old," "maybe old," "don't know," "maybe new," and "sure new"). In addition, as a baseline measure to assess participants' ability to guess the location of objects within a scene in the absence of any memory, we also plotted heatmaps for new (i.e., nonstudied) scenes (Figure 2A). The heatmaps of new scenes were based on the recalled locations relative to the object locations for those same scenes in the opposite counterbalance, in which participants had seen the scene at study; a new scene's designation as congruent or incongruent was therefore determined by the condition in which it was presented in the opposite counterbalance.

Table 1

Recognition Response Proportions for Each Scene Type

	"Sure New"	"Maybe New"	"Don't Know"	"Maybe Old"	"Sure Old"	"Recollect Old"
Old Scenes	7%	10%	11%	16%	24%	32%
New Scenes	37%	28%	14%	12%	5%	3%

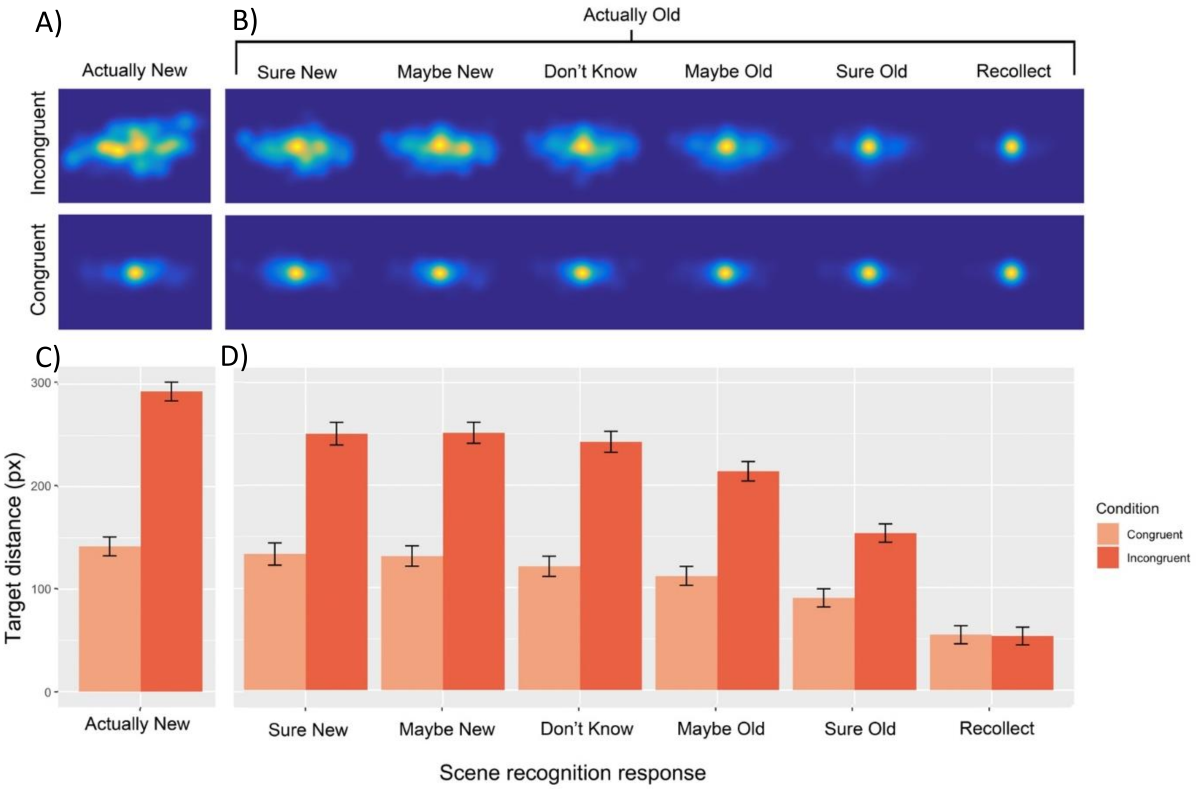


Figure 2. Object location memory for schema-congruent and incongruent target objects in new scenes (A) and old scenes (B). Each heat map illustrates the distribution of recalled locations for the scenes, normalized such that the center of the heatmap represents the location of the target object. Thus, heatmaps tightly focused on the center-point —as in the recollected scenes— indicate high spatial accuracy, whereas more distributed heatmaps indicate poorer spatial accuracy. C and D) Spatial memory accuracy measured as the distance between the recalled location and the studied object location. Higher values indicate lower accuracy. The least-squares means derived from the models used in the analyses are plotted, and the error bars represent the standard error of these estimated means from the model.

An examination of Figure 2B shows that spatial memory was generally better for congruent than incongruent scenes (i.e., more tightly focused recall heatmaps for schema-congruent than incongruent scenes). Importantly, this congruency effect was dependent on the type of memory participants had for the scenes: Unconscious memory reduced the congruency effect compared to new scenes ("Sure new" old scenes versus new scenes), and the congruency effect weakened further with increasing familiarity strength ("Don't know" through "Sure old" scenes). Perhaps surprisingly, for recollected scenes, the congruency effect was eliminated entirely such that spatial accuracy was similar for schema-congruent and incongruent scenes; in fact, in contrast to every other response type, performance was numerically better in incongruent than congruent recollected scenes.

These effects were analyzed by examining the distance between the recalled location and the actual target location (i.e., “target distance”; Figure 2C-D). For recollected scenes, target distance did not differ between schema-congruent and incongruent scenes, *t*(10271)=0.23, *p*=.82, *d*=0.004, indicating that schema congruency had no significant effect on spatial memory for scenes that were recollected. Analysis using Bayes factors provided substantial evidence for the null hypothesis (BF_{10}=0.16)[8] that there was no schema congruency effect within recollected scenes. Moreover, a direct comparison of performance in the recollected scenes and the high-confidence familiar scenes (i.e., “sure old” scenes) revealed a significant interaction such that the effect of schema congruency on target distance was larger for scenes recognized on the basis of familiarity compared to those judged to be recollected, *t*(4292)=-7.68, *p*<.0001, *d*= -0.23.

To probe the relationship between episodic familiarity strength and schema knowledge, we examined whether the linear gradient of familiarity-based responses (i.e., “sure new” through “sure old” responses) interacted with congruency to predict target distance. We found that as familiarity strength increased, the effect of schema congruency on target distance decreased, β=-.12, *t*(5242)=-4.97, *p*<.0001. These results indicate that the schema congruency effects were present in familiar scenes, but were weaker for more familiar items.

To examine the impact of schema congruency under conditions in which there was no contribution of memory, we examined target distance for new, non-studied scenes, and found that spatial accuracy was significantly greater for schema-congruent than incongruent scenes,

---

[8] 1 By convention, a BF_{10} < 0.33 indicates substantial evidence for the null hypothesis (Jeffreys, 1961).

t(2551)=21.21, *p*<.0001, *d*=0.84. Moreover, a direct comparison of old scenes that could only have unconscious memory (i.e., old scenes receiving a “sure new” response) with truly new scenes revealed that there was a significant interaction such that the schema congruency effect was larger for new scenes compared to unconsciously recognized old scenes, *t*(3096)=-2.84, *p*=.005, *d*=-0.1. This indicates that even in the absence of conscious memory, schema congruency effects on spatial memory are reduced by unconscious memory for a previously viewed scene. Overall, the results thus far indicate that schema congruency has its largest effects on spatial accuracy when memory for the scenes is weakest.

To better characterize the effects of schema congruency on spatial memory accuracy, we further examined the recalled locations in schema-incongruent scenes across all memory responses. We found that when participants made a spatial location error, they did not select a location randomly, but rather they often incorrectly recalled that the object had been in the schema-congruent location. For example, Figure 3A-B shows the distribution of recalled locations for one of the scenes when the object was in the schema-congruent or schema-incongruent location, respectively. For the congruent condition, the average recalled location was quite close to the target location. For the incongruent condition, most recalled locations were close to the target location, but there was also a tendency to falsely recall that the object had been in the congruent location. To quantify this latter effect, we calculated a new target distance measure for each incongruent scene by measuring the distance between the clicked location in the test phase and the target location from the *congruent* version of that scene (i.e., *lure target distance*). Importantly, the included participants had never actually seen the congruent version of these scenes in the study phase. As a baseline against which to compare these lure target distance values, we also calculated target distance values between the clicked locations and randomly selected target locations from other scenes, and took the average over 100 iterations of this process. Lure target distance values that are lower than the random baseline would imply above-chance likelihood of clicking near the lure location; that is, participants would be more likely to click near the lure location than any other incorrect location. This would indicate a semantic bias towards the congruent—but incorrect—location.

For the analysis, we only used trials in which participants did not successfully click within 25px of the incongruent location (but note that all results outlined below were significant to *p*<.001 when all trials were used). We found that participants were more likely than chance to click near the lure, congruent location in new scenes, *p*<.0001, and in all old scenes that were not recollected, *p*s<.005 (Figure 3C). This indicates that in new scenes, and in old scenes that were forgotten or familiar, participants' spatial accuracy was reduced in part due to semantic bias towards the schema-congruent location. In contrast, in recollected scenes, participants were less likely than chance to falsely recall the lure congruent location, *t*(923)=-4.09, *p*<.0001, *d*=-.27. This suggests that even when spatial memory failed in recollected scenes (i.e., was >25px from the target location), participants were able to recollect that a target was not in a congruent location—thus sparing them from incorrect semantic bias. In other words, participants' recollection may have eliminated congruent locations as possible locations, thus improving their performance even when they did not remember the incongruent target location per se. Together, these results indicate not only that schema knowledge actively undermines spatial memory performance in incongruent scenes, but that strong episodic memory—particularly recollection—protects against this effect.

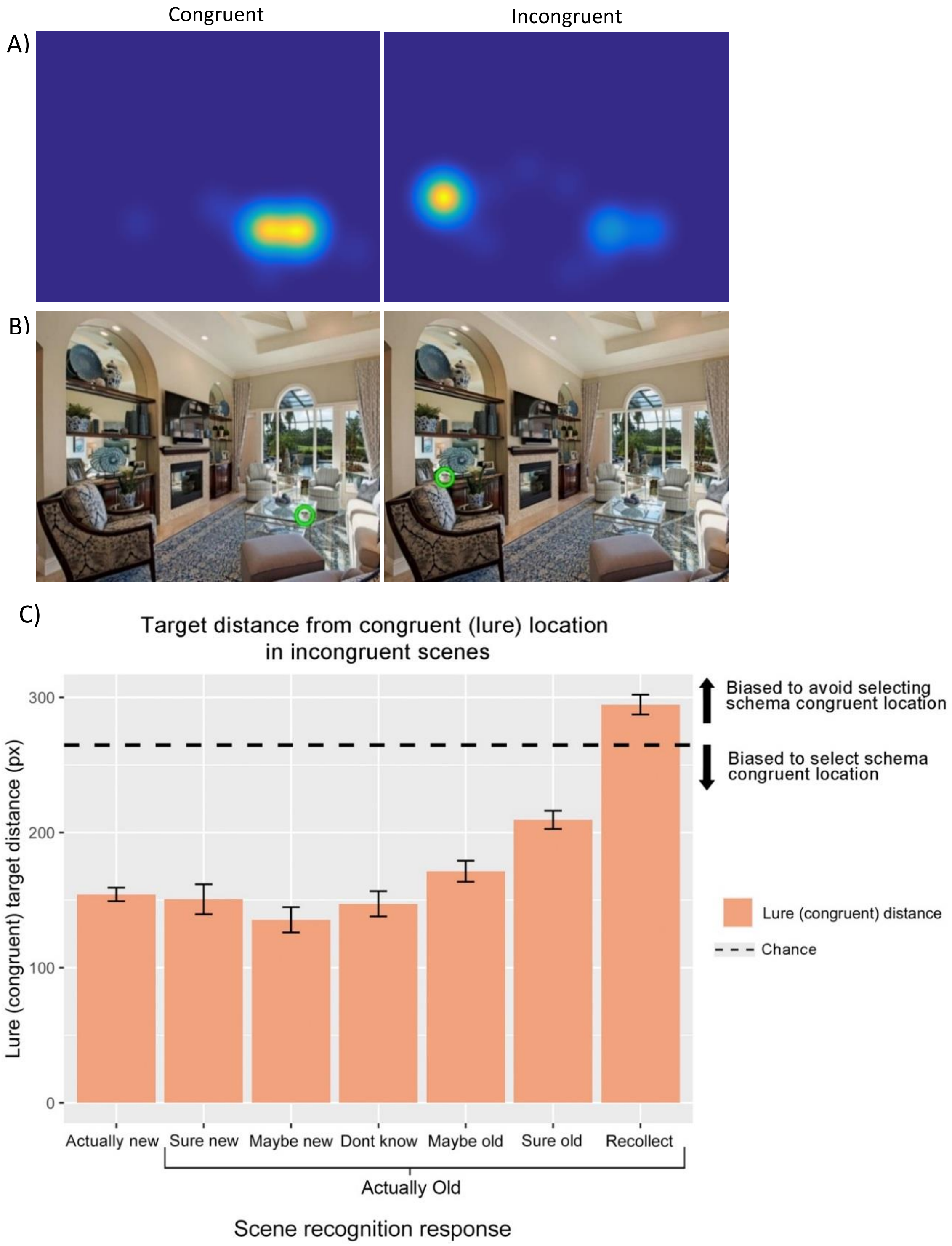


Figure 3. The effect of schema knowledge on spatial recall errors for incongruent scenes. A) Example of the effect of schema congruency on spatial memory. The heatmaps are smoothed aggregate density maps of the click locations made on the congruent and incongruent version of the same scene. For example, the congruent heatmap includes the test phase click location from each participant who saw the congruent version of the scene. The incongruent heatmap suggests

that many of the errors in these trials were due to participants erroneously choosing the congruent region. B) The congruent and incongruent versions of the scene. The target is circled in green for each case. C) Distance between the recalled object location and the schema-congruent location for objects studied in an incongruent location, plotted for each type of recognition response. Trials that were correctly recalled (i.e., <25 pixels from the studied location) have been excluded. The dashed line represents chance performance, which is the average distance between the recalled locations and a randomly selected target location. Values below the chance line indicate that participants' selected locations were more likely than chance to be near the schema-congruent region, and indicate that errors on those incongruent scenes may be driven by schema bias. Values above the chance line, as in the recollect responses, indicate that the selected locations were less likely than chance to be near a schema congruent region; this suggests that errors in these trials were not driven by schema bias. Least-squares means derived from the models used in the analyses are plotted, and the error bars represent the standard error of these estimated means from the model.

Although the results of the current study were quite robust such that the interactions between schema congruence and episodic memory were highly significant, we wished to determine whether the results would replicate, so we conducted a second experiment with a different group of participants using the same method (Figure 4). Importantly, the pattern of results was found to replicate in this second experiment (compare Figure 2 to Figure 4), and all of the significant results reported above replicated. Specifically, the effect of congruency was significantly reduced due to unconscious memory (i.e., in "sure new" old scenes compared to new scenes), *t*(1315)=-2.44, *p*=.015, *d*=-0.14. Schema congruency effects also decreased as familiarity strength increased, β=-.16, *t*(2110)=-4.10, *p*<.0001. Furthermore, there was no congruency effect in recollected scenes, *t*(4455)=-0.32, *p*=.75, *d*=-0.01, and recollection uniquely protected from bias towards congruent regions, *t*(382)=2.66, *p*=.008, *d*=0.27.

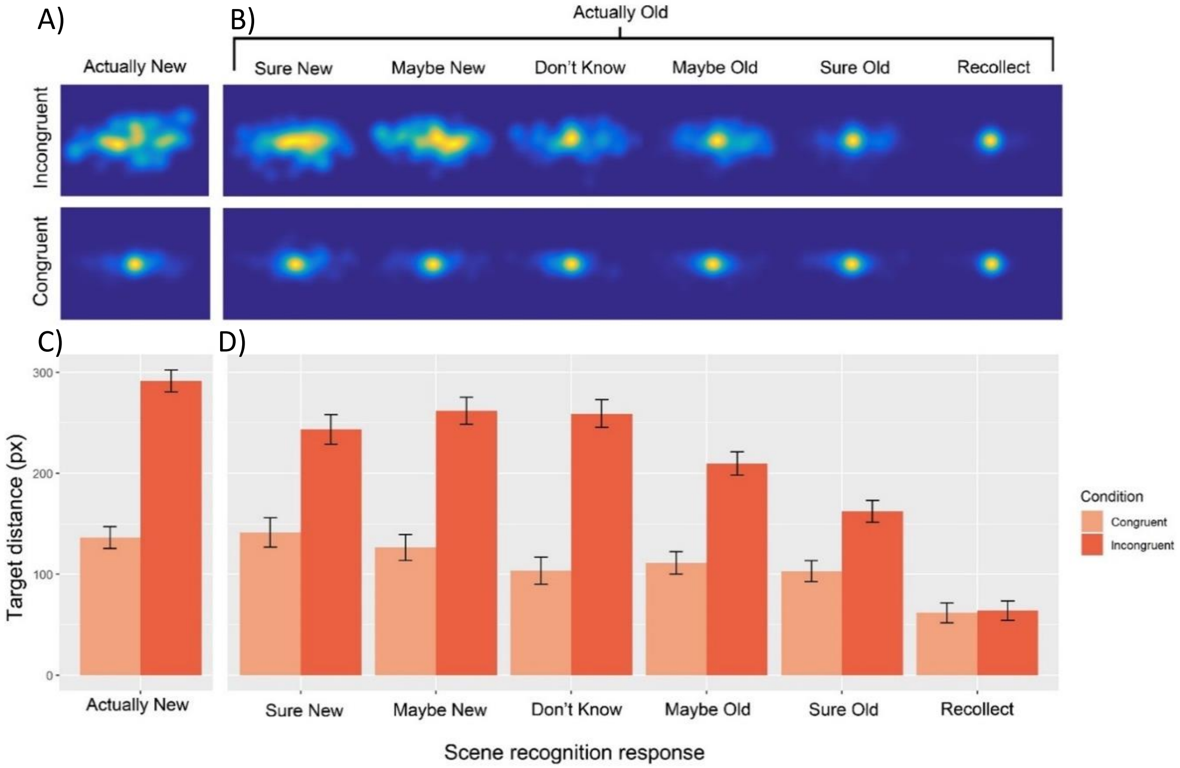


Figure 4. Replication experiment data, showing object location memory for schema-congruent and incongruent target objects in new scenes (A) and old scenes (B). Each heat map illustrates the distribution of recalled locations for the scenes, normalized such that the center of the heatmap represents the location of the target object. Thus, heatmaps tightly focused on the center-point —as in the recollect responses—indicate high spatial accuracy, whereas more distributed heatmaps indicate poorer spatial accuracy. C and D) Spatial memory accuracy measured as the distance between the recalled location and the studied object location. The least-squares means derived from the models used in the analyses are plotted, and the error bars represent the standard error of these estimated means from the model.

Discussion

In the present study, we examined how schema knowledge interacts with different recognition memory processes to influence spatial memory accuracy. Participants first searched scenes for a target object that was in either a schema-congruent or a schema-incongruent location. In a subsequent spatial recall phase, participants indicated where in each scene they remembered the target object having been located in the earlier search task. In addition, they gave a confidence-based recognition memory judgment for each scene that allowed us to isolate recollection, familiarity, and unconscious memory. We found a robust schema congruency effect such that spatial recall was more accurate for objects in congruent than incongruent locations. Importantly, the magnitude of this effect decreased as memory strength for the scenes increased. That is, the congruency advantage was largest for new scenes, decreased with both unconscious memory and familiarity confidence, and was absent entirely for recollected scenes. In addition, poorer spatial memory performance in incongruent scenes reflected a tendency to falsely recall the target as having been in a schema-congruent location, but only in scenes that were not recollected: When recollection occurred, participants were significantly biased against selecting the congruent locations—even when they did not remember the correct incongruent location—suggesting that recollection was able to oppose the effects of schema bias.

The present results provide strong support for neural models proposing competitive interactions between episodic memory and schema knowledge (Gilboa & Marlatte, 2017; Sommer, 2017; van Kesteren et al., 2012; Wagner et al., 2015). These models suggest that during learning of arbitrary object-location associations, the episodic system—which is reliant on the hippocampus—is critical for binding objects to specific locations. However, when object locations are consistent with established spatial schemas, learning can be supported by a separate schema-learning system, which is reliant on the cortex and the ventromedial prefrontal cortex (vmPFC). These episodic memory and schema systems are assumed to be competitive and mutually inhibitory, as suggested by studies showing that learning schema-related knowledge is related to increased cortical and vmPFC activation and reduced hippocampal activity (for review see Gilboa & Marlatte, 2017). However, direct behavioral evidence that these systems compete at retrieval has been lacking. To our knowledge, the current results are the first to provide such direct evidence.

The reduction of schema effects by memory was strongest for scenes that were recollected, but was also evident in familiar scenes such that the effect of schema congruency decreased as familiarity confidence increased. This indicates that both recollection and familiarity can compete with and effectively oppose the effects of schema knowledge. The competitive interaction models have not differentiated between recollection and familiarity-based memory, so the current results present new challenges to those models and suggest areas for further investigation. For example, the hippocampus plays a critical role in supporting recollection (Bastin et al., 2019; Eichenbaum et al., 2007 but see Wais et al., 2006), and therefore our finding that recollection competes with schema knowledge supports the claim that the hippocampus and regions supporting schema knowledge are mutually inhibitory (Gilboa & Marlatte, 2017; van Kesteren et al., 2012). However, medial temporal lobe regions outside the hippocampus such as the perirhinal cortex support familiarity (Aggleton & Brown, 1999; Eichenbaum et al., 2007). Thus, the current results suggest that schema-related regions may also be in direct competition with regions such as the perirhinal cortex, which has not yet been considered.

The extent to which schemas and memory trade off may depend upon the precision of spatial information provided by memory. That is, prior work has indicated that recollection provides high-precision spatial information, whereas familiarity provides low-precision, gist-like spatial information (Kolarik et al., 2016, 2018). Object location predictions provided by schema knowledge, on the other hand, are typically diffuse and probabilistic rather than precise (Biederman, 1981; Torralba et al., 2006). Therefore, when high-precision recollection is available, predictions provided by recollection-related representations may be prioritized over schema predictions. Similarly, because familiarity and presumably unconscious memory provide some degree of precision that is above and beyond the probabilistic spatial information supported by schema knowledge, schematic spatial predictions may be down-weighted in favor of these more deterministic memory predictions. If this account is correct, it should be possible to modify the outcome of schema-memory competition by varying the relative precision of the spatial information supported by each. For example, schema knowledge may be more heavily weighted if it supports precise spatial information (e.g., the location of a door handle vs. the location of a cellphone), whereas recollection may be more heavily weighted if it provides more precise information (e.g., when I attend to where I placed my phone vs. attend to other aspects of the episode).

The schema congruency effects observed in the current study were smaller for unconsciously recognized scenes than for truly new scenes. That is, for studied scenes that participants were highly confident had not been studied, participants had better spatial recall than if the scenes had not been studied at all. These results suggest that even unconscious forms of memory can support memory for object locations. Interestingly, prior studies finding unconscious memory effects have typically used implicit measures such as eye movements and search speed (see Hannula & Greene, 2012), but the present study points to a possible role of unconscious memory even in explicit spatial recall decisions. Moreover, combining the present paradigm with neuroimaging methods or patient populations may be particularly useful for informing the debates surrounding hippocampal involvement in unconscious memory for associations (Hannula & Greene, 2012).

In sum, it is clear that memory and schema knowledge simultaneously contribute to a wide variety of everyday behaviors, but how they are resolved to influence behavior has been a subject of debate. Although schemas and memory can synergistically enhance performance in some cases, schemas can also lead to systematic errors when they are inconsistent with individual past experiences. The current results support recent neurocognitive models proposing that this conflict is resolved through competition, and show that the effects of schema knowledge can be effectively eliminated when memory provides strong episodic information. These findings of memory-dependent schema suppression not only point to ways to improve models of schema-memory interactions, but have implications for schema research in a variety of other areas.

References

Aggleton, J. P., & Brown, M. W. (1999). Episodic memory, amnesia, and the hippocampal-anterior thalamic axis. In *Behavioral and Brain Sciences* (Vol. 22, Issue 3, pp. 425–444). Cambridge University Press. https://doi.org/10.1017/S0140525X99002034

Alba, J. W., & Hasher, L. (1983). Is memory schematic? *Psychological Bulletin*, *93*(2), 203–231. https://doi.org/10.1037/0033-2909.93.2.203

Anderson, J. R. (1981). Effects of prior knowledge on memory for new information. *Memory & Cognition*, *9*(3), 237–246. https://doi.org/10.3758/BF03196958

Bartlett, F. C., & Burt, C. (1932). Remembering: a Study in Experimental and Social Psychology. *British Journal of Educational Psychology*, *3*(2), 187–192. https://doi.org/10.1111/j.2044-8279.1933.tb02913.x

Bastin, C., Besson, G., Simon, J., Delhaye, E., Geurten, M., Willems, S., & Salmon, E. (2019). An Integrative Memory model of recollection and familiarity to understand memory deficits. *Behavioral and Brain Sciences*, *42*. https://doi.org/10.1017/S0140525X19000621

Bein, O., Reggev, N., & Maril, A. (2014). Prior knowledge influences on hippocampus and medial prefrontal cortex interactions in subsequent memory. *Neuropsychologia*, *64*, 320–330. https://doi.org/10.1016/j.neuropsychologia.2014.09.046

Biederman, I. (1981). On the Semantics of a Glance at a Scene. In *Perceptual Organization* (pp. 213–253). Routledge. https://doi.org/10.4324/9781315512372-8

Boettcher, S. E. P., Draschkow, D., Dienhart, E., & Võ, M. L.-H. (2018). Anchoring visual search in scenes: Assessing the role of anchor objects on eye movements during visual search. *Journal of Vision*, *18*(13), 11. https://doi.org/10.1167/18.13.11

Cox, W. T. L., Abramson, L. Y., Devine, P. G., & Hollon, S. D. (2012). Stereotypes, Prejudice, and Depression: The Integrated Perspective. *Perspectives on Psychological Science*, *7*(5), 427–449. https://doi.org/10.1177/1745691612455204

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

Dismukes, R. K. (2008). Prospective Memory in Aviation and Everyday Settings. *Prospective Memory: Cognitive, Neuroscience, Developmental, and Applied Perspectives*, 411–428.

Eichenbaum, H., Yonelinas, A. P., & Ranganath, C. (2007). The medial temporal lobe and recognition memory. *Annual Review of Neuroscience*, *30*, 123–152. https://doi.org/10.1146/annurev.neuro.30.051606.094328

Ghosh, V. E., & Gilboa, A. (2014). What is a memory schema? A historical perspective on current neuroscience literature. *Neuropsychologia*, *53*(1), 104–114. https://doi.org/10.1016/j.neuropsychologia.2013.11.010

Gilboa, A., & Marlatte, H. (2017). Neurobiology of Schemas and Schema-Mediated Memory. *Trends in Cognitive Sciences*, *21*(8), 618–631. https://doi.org/10.1016/j.tics.2017.04.013

Greve, A., Cooper, E., Tibon, R., & Henson, R. N. (2019). Knowledge is power: Prior knowledge aids memory for both congruent and incongruent events, but in different ways. *Journal of Experimental Psychology: General*, *148*(2), 325–341. https://doi.org/10.1037/xge0000498

Hannula, D. E., & Greene, A. J. (2012). The hippocampus reevaluated in unconscious learning and memory: at a tipping point? *Frontiers in Human Neuroscience*, *6*, 80. https://doi.org/10.3389/fnhum.2012.00080

Hawke, L. D., & Provencher, M. D. (2011). Schema Theory and Schema Therapy in Mood and

Anxiety Disorders: A review. *Journal of Cognitive Psychotherapy*, *25*(4), 257–276. https://doi.org/10.1891/0889-8391.25.4.257

Henderson, J. M. (2003). Human gaze control during real-world scene perception. *Trends in Cognitive Sciences*, *7*(11), 498–504. https://doi.org/10.1016/j.tics.2003.09.006

Hole, G. (2014). *The Psychology of Driving*. Psychology Press. https://books.google.com/books?hl=en&lr=&id=epX_AwAAQBAJ&oi=fnd&pg=PP1&ots=C6QYRNRUiO&sig=33beW8FjPlnZS2YvWkYXT_b3mRY

Johnson, M. K. (1997). Source monitoring and memory distortion. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *352*(1362), 1733–1745. https://doi.org/10.1098/rstb.1997.0156

Kolarik, B. S., Baer, T., Shahlaie, K., Yonelinas, A. P., & Ekstrom, A. D. (2018). Close but no cigar: Spatial precision deficits following medial temporal lobe lesions provide novel insight into theoretical models of navigation and memory. *Hippocampus*, *28*(1), 31–41. https://doi.org/10.1002/hipo.22801

Kolarik, B. S., Shahlaie, K., Hassan, A., Borders, A. A., Kaufman, K. C., Gurkoff, G., Yonelinas, A. P., & Ekstrom, A. D. (2016). Impairments in precision, rather than spatial strategy, characterize performance on the virtual Morris Water Maze: A case study. *Neuropsychologia*, *80*, 90–101. https://doi.org/10.1016/j.neuropsychologia.2015.11.013

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest Package: Tests in Linear Mixed Effects Models. *Journal of Statistical Software*, *82*(13), 1–19. https://doi.org/10.18637/jss.v082.i13

Lew, A. R., & Howe, M. L. (2017). Out of Place, Out of Mind: Schema-Driven False Memory Effects for Object-Location Bindings. *Journal of Experimental Psychology: Learning,* *Memory, and Cognition*, *43*(3), 404–421. https://doi.org/10.1037/xlm0000317.supp

McClelland, J. L. (2013). Incorporating rapid neocortical learning of new schema-consistent information into complementary learning systems theory. *Journal of Experimental Psychology: General*, *142*(4), 1190–1210. https://doi.org/10.1037/a0033812

McClelland, J. L., McNaughton, B. L., & O'Reilly, R. C. (1995). Why there are complementary learning systems in the hippocampus and neocortex: Insights from the successes and failures of connectionist models of learning and memory. *Psychological Review*, *102*(3), 419–457. https://doi.org/10.1037/0033-295X.102.3.419

McKenzie, S., Frank, A. J., Kinsky, N. R., Porter, B., Rivière, P. D., & Eichenbaum, H. (2014). Hippocampal representation of related and opposing memories develop within distinct, hierarchically organized neural schemas. *Neuron*, *83*(1), 202–215. https://doi.org/10.1016/j.neuron.2014.05.019

McKenzie, S., Robinson, N. T. M., Herrera, L., Churchill, J. C., & Eichenbaum, H. (2013). Learning causes reorganization of neuronal firing patterns to represent related experiences within a hippocampal schema. *Journal of Neuroscience*, *33*(25), 10243–10256. https://doi.org/10.1523/JNEUROSCI.0879-13.2013

Plant, K. L., & Stanton, N. A. (2012). Why did the pilots shut down the wrong engine? Explaining errors in context using Schema Theory and the Perceptual Cycle Model. *Safety Science*, *50*(2), 300–315. https://doi.org/10.1016/j.ssci.2011.09.005

Preston, A. R., & Eichenbaum, H. (2013). Interplay of hippocampus and prefrontal cortex in memory. *Current Biology*, *23*(17), R764–R773. https://doi.org/10.1016/j.cub.2013.05.041

Ramey, M. M., Henderson, J. M., & Yonelinas, A. P. (2020). The Spatial Distribution of Attention Predicts Familiarity Strength During Encoding and Retrieval. *Journal of*

Experimental Psychology: General, *149*(11), 2046–2062. https://doi.org/10.1037/xge0000758

Ramey, M. M., Yonelinas, A. P., & Henderson, J. M. (2019). Conscious and unconscious memory differentially impact attention: Eye movements, visual search, and recognition processes. *Cognition*, *185*, 71–82. https://doi.org/10.1016/J.COGNITION.2019.01.007

Ramey, M. M., Yonelinas, A. P., & Henderson, J. M. (2020). Why do we retrace our visual steps? Semantic and episodic memory in gaze reinstatement. *Learning and Memory, 27*(7), 275–283. https://doi.org/10.1101/lm.051227.119

Robin, J., & Moscovitch, M. (2017). Details, gist and schema: hippocampal–neocortical interactions underlying recent and remote episodic and spatial memory. *Current Opinion in Behavioral Sciences*, *17*, 114–123. https://doi.org/10.1016/j.cobeha.2017.07.016

Roediger, H. L., & McDermott, K. B. (1995). Creating false memories: Remembering words not presented in lists. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *21*(4), 803–814. https://doi.org/10.1037/0278-7393.21.4.803

Rumelhart, D. E. (2017). Schemata: The building blocks of cognition. *Theoretical Issues in Reading Comprehension: Perspectives from Cognitive Psychology, Linguistics, Artificial Intelligence and Education*, 33–58. https://doi.org/10.4324/9781315107493-4

Sommer, T. (2017). The Emergence of Knowledge and How it Supports the Memory for Novel Related Information. *Cerebral Cortex*, *27*(3), 1906–1921. https://doi.org/10.1093/cercor/bhw031

Sweegers, C. C. G., Coleman, G. A., van Poppel, E. A. M., Cox, R., & Talamini, L. M. (2015). Mental schemas hamper memory storage of goal-irrelevant information. *Frontiers in Human Neuroscience*, *9*, 629. https://doi.org/10.3389/fnhum.2015.00629

Torralba, A., Oliva, A., Castelhano, M. S., & Henderson, J. M. (2006). Contextual guidance of eye movements and attention in real-world scenes: The role of global features in object search. *Psychological Review*, *113*(4), 766–786. https://doi.org/10.1037/0033-295X.113.4.766

Tulving, E. (1982). Synergistic ecphory in recall and recognition. *Canadian Journal of Psychology/Revue Canadienne de Psychologie*, *36*(2), 130–147. https://doi.org/10.1037/h0080641

Tulving, E. (1985). Memory and consciousness. *Canadian Psychology/Psychologie Canadienne*, *26*(1), 1–12. https://doi.org/10.1037/h0080017

van Buuren, M., Kroes, M. C. W., Wagner, I. C., Genzel, L., Morris, R. G. M., & Fernández, G. (2014). Initial investigation of the effects of an experimentally learned schema on spatial associative memory in humans. *Journal of Neuroscience*, *34*(50), 16662–16670. https://doi.org/10.1523/JNEUROSCI.2365-14.2014

van Kesteren, M. T. R., Beul, S. F., Takashima, A., Henson, R. N., Ruiter, D. J., & Fernández, G. (2013). Differential roles for medial prefrontal and medial temporal cortices in schema-dependent encoding: From congruent to incongruent. *Neuropsychologia*, *51*(12), 2352–2359. https://doi.org/10.1016/j.neuropsychologia.2013.05.027

van Kesteren, M. T. R., Rijpkema, M., Ruiter, D. J., & Fernández, G. (2010). Retrieval of associative information congruent with prior knowledge is related to increased medial prefrontal activity and connectivity. *Journal of Neuroscience*, *30*(47), 15888–15894. https://doi.org/10.1523/JNEUROSCI.2674-10.2010

van Kesteren, M. T. R., Ruiter, D. J., Fernández, G., & Henson, R. N. (2012). How schema and novelty augment memory formation. *Trends in Neurosciences*, *35*(4), 211–219.

https://doi.org/10.1016/j.tins.2012.02.001

Võ, M. L.-H., Boettcher, S. E., & Draschkow, D. (2019). Reading scenes: how scene grammar guides attention and aids perception in real-world environments. In *Current Opinion in Psychology* (Vol. 29, pp. 205–210). Elsevier B.V. https://doi.org/10.1016/j.copsyc.2019.03.009

Võ, M. L.-H., & Wolfe, J. M. (2013). The interplay of episodic and semantic memory in guiding repeated search in scenes. *Cognition*, *126*(2), 198–212. https://doi.org/10.1016/j.cognition.2012.09.017

Wagner, I. C., van Buuren, M., Kroes, M. C. W., Gutteling, T. P., van der Linden, M., Morris, R. G., & Fernández, G. (2015). Schematic memory components converge within angular gyrus during retrieval. *ELife*, *4*, e09668. https://doi.org/10.7554/eLife.09668.001

Wais, P. E., Wixted, J. T., Hopkins, R. O., & Squire, L. R. (2006). The hippocampus supports both the recollection and the familiarity components of recognition memory. *Neuron*, *49*(3), 459–466. https://doi.org/10.1016/j.neuron.2005.12.020

Wixted, J. T. (2007). Dual-process theory and signal-detection theory of recognition memory. *Psychological Review*, *114*(1), 152–176. https://doi.org/10.1037/0033-295X.114.1.152

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