Associations among attentional state, retrieval quality, and mnemonic discrimination

Christopher N. Wahlheim [a,*], Sydney M. Garlitch [b], Rawan M. Mohamed [a], Blaire J. Weidler [c]

[a] Department of Psychology, University of North Carolina at Greensboro, United States
[b] Department of Behavioral Sciences, Millikin University, United States
[c] Department of Psychology, Towson University, United States

ABSTRACT

Memory specificity is shown when participants reject lures that are similar to studied objects. Lure rejections may reflect hippocampal pattern separation that encodes objects distinctively. However, lure features shared with studied objects may evoke pattern completion of varying quality. This was shown when self-reported attention during study promoted lure rejections and false alarms. We used an experimental and individual differences approach to examine the roles of attentive encoding and retrieval quality in lure classifications. An object-based mnemonic discrimination task included thought probes during study and subjective retrieval reports after recognition responses. On-task reports reflecting attentive encoding were associated with lure rejections and false alarms within-and between-subjects. Additionally, accurate lure and target classifications were more strongly associated with subjective recollection following on- than off-task reports. Collectively, these results suggest that attention during study was associated with recollection of criterial features that differentiated existing memories from perceptual inputs.

Introduction

People behave according to routines and habits. They visit the same coffee shops, talk with the same work colleagues, and walk their dogs on the same routes. New experiences often repeat aspects of the past—but they are never identical—as when a person notices that their favorite coffee shop changed its seasonal menu items. To maintain updated memories for recent events, people must initially distinguish those events from memories of related events. Doing so requires attending to event details to notice relationships between events, but people vary widely in how they allocate their attention to the environment. The ability to distinguish memories from perceptions may thus vary with how people attend to event details during encoding. We identified these relationships here by examining how attention while encoding events associates with recognition of similar-but-not-identical events. We also identified the role of self-reported recollection of related memories in such mnemonic discrimination. The findings we report below are relevant to literatures on the processes underlying mnemonic discrimination and individual differences in the relationship between attention and memory. Before describing our specific approach, we summarize key

background issues from those two literatures in turn.

Processes underlying mnemonic discrimination

The human memory system encodes similar events distinctively via hippocampal pattern separation. Pattern separation is a computational process whereby perceptual inputs are encoded orthogonally from similar memories (Marr, 1971). The mnemonic discrimination that results from pattern separation prevents catastrophic memory interference between similar events (McClelland et al., 1995; Norman & O'Reilly, 2003). This process could support recalling, for example, recent menu items as being different from previous items. Pattern separation is often assessed using mnemonic discrimination tasks that require rejecting lures that are similar but not identical to earlier events. However, behavioral tasks cannot purely measure pattern separation when features of similar events cue retrievals, a process called pattern completion (cf. Marr, 1971). Behavioral measures of pattern separation may thus be susceptible to contamination from pattern completion when overlapping features are effectively attended to and encoded.

In mnemonic discrimination tasks, participants study memoranda,

such as pictures of objects or scenes, and then complete a modified recognition test. As in typical recognition tasks, the test often includes targets that are exact repetitions of studied items and new foils that did not appear during study. Unlike typical recognition tasks, the test also includes items of the same identity as studied items but with different features (i.e., similar lures). For example, a participant could study a truck, then be tested on another truck with similar coloring and positioned in a different orientation. Pattern separation is assumed to contribute to the rejection of lures when participants identify them as "similar." However, pattern completion may also contribute when people can recollect studied item features (Stark et al., 2019). This assertion echoes the long-standing view that no task is process pure (e.g., Jacoby, 1991). The contribution of pattern completion to lure rejections has been shown routinely in false memory studies indicating that recollecting criterial features allows participants to detect and reject lures (Bowman & Dennis, 2016; Gallo, 2004; Lampinen et al., 2006). Moreover, although "similar" classifications can reflect recollection-based rejections, participants can also use these responses to indicate low-confidence memories for studied items (Loiotile & Courtney, 2015). Collectively, these findings highlight the caveats of using mnemonic discrimination tasks to measure pattern separation.

The literature on neural approaches of assessing pattern separation also highlights the complexity of separately assessing pattern separation and completion. Researchers have argued that lure rejections reflect pattern separation because rejections vary with hippocampal function and structure (for a review, see Stark et al., 2019). However, neurocomputational models propose the hippocampus also supports pattern completion (e.g., Marr, 1971; Rolls, 2013). This mixture of pattern separation and completion may involve differential participation of hippocampal subfields, primarily the dentate gyrus (DG) and CA3 (for reviews, see Rolls, 2016; Yassa & Stark, 2011). Because the hippocampus supports both processes, researchers have debated whether these processes operate in parallel or compete for resources (Holden & Gilbert, 2012; Hunsaker & Kesner, 2013; Nakashiba et al., 2012; Ngo et al., 2021; Yassa & Stark, 2011). This debate has implications for the inferences about the extent to which pattern separation and completion contribute to lure classification patterns in mnemonic discrimination tasks.

Researchers have examined the roles of pattern separation and completion in these tasks by comparing lure classification patterns across groups varying in hippocampal function. A key theoretical assumption is that lure rejections should be inversely related to false alarms to the extent that pattern separation and completion compete for hippocampal resources. Accordingly, groups with poorer hippocampal function should show decreased lure rejections and increased lure false alarms compared to controls. When classifying similar lures, healthy older adults, who have poorer hippocampal function than their younger counterparts, show fewer rejections and more false alarms (e.g., Davidson et al., 2019; Stark et al., 2013; Wahlheim et al., 2022). Of relevance to the present study, computational modeling suggests that this age-related difference reflects older adults' attention deficit leading to poorer encoding of studied items (Huffman & Stark, 2017). However, such group differences are not always observed. For example, reduced lure rejections have been observed without increased false alarms in patients with Alzheimer's disease (Ally et al., 2013) and hippocampal damage (Kirwan et al., 2012). Collectively, these findings suggest that pattern separation and completion may not necessarily tradeoff and that encoding variability resulting from attentional states played a key role in lure classifications.

Eye tracking methods have also been used to examine the role of encoding variability during study in the contributions of these processes to lure classifications. Although the present study does not incorporate such measures, findings from those studies provide a broader context for understanding associations between attention during study and lure classifications. Those studies assumed that more fixations during study indicated better encoding. Similar to the aforementioned modeling outcome suggesting that poorer encoding leads to more lure false alarms and fewer rejections, false alarms have been shown in association with fewer study fixations than correct rejections (Bjornn et al., 2022; Molitor et al., 2014). Also, lure rejections and false alarms corresponded with more study fixations than items leading to target recognition, implicating pattern completion in lure classifications. This role of pattern completion was also supported by fixation differences in a two-alternative forced-choice task (Rollins et al., 2019). False alarms to lures paired with targets were associated with more study fixations to objects corresponding with lures. This suggests that more attentive encoding led to more inaccurate memory-based classifications. Collectively, these findings suggest that more attention to studied items promotes the use of pattern completion during lure classifications. However, although the accuracy of such lure classifications may be improved by better encoding, that is not always the case.

Other methods have also implicated a role for encoding variability in patterns of lure classifications. Improving the encoding of study items via repetition has been claimed to impair bias-corrected lure discrimination by decontextualizing memories (Reagh & Yassa, 2014). However, studies that appropriately accounted for repetition-induced encoding differences showed that repetitions actually improved such discrimination (Loiotile & Courtney, 2015; Zhang & Hupbach, 2020), even when added contextual variability could have accelerated item decontextualization (Zhang & Hupbach, 2023). Together, these findings suggest that improving the encoding of studied items increased the use of pattern completion to reject corresponding lures. Relatedly, we have examined the role encoding variability in lure classifications using a measure of self-reported attentional states during study (Wahlheim et al., 2023). While participants encoded study items, they intermittently reported their task engagement as being on– or off-task. We assumed that on-task reports reflected attentional engagement and better encoding. Similar to the studies above, we showed that on-task reports were associated with more lure rejections. However, we also showed that on-task reports were associated with more lure false alarms. These findings suggested that effective encoding promoted pattern completion that supported lure rejection, but only when retrieved memories included distinguishing details.

Although memory-based lure rejections may depend on encoding, the associations between encoding variability and pattern completion in lure classifications have not been well-characterized. This was the primary goal of the present study. However, two studies (Kim & Yassa, 2013; Szőllősi et al., 2020) have provided initial characterizations of the contributions of retrieval processes to lure classifications using variants of the remember/know procedure (Tulving, 1985). In those studies, after lure rejections and false alarms, participants classified their retrieval states as "remember" when they *recollected* studied objects and "know" when their memory for an object felt *familiar* but they could not recollect the object. The first study showed that lure rejections and false alarms were associated with more "remember" than "know" responses (Kim & Yassa, 2013). However, the second study showed that lure false alarms were associated with more "remember" than "know" responses, while lure rejections were associated with more "know" than "remember" responses (Szőllősi et al., 2020). This differential use of retrieval processes may reflect a longer (10 min) study-test interval and/or "guess" response option in Szőllősi et al. (2020). More generally, these mixed findings suggest that mnemonic variables determine the extent to which recollection and familiarity support lure classifications. However, the moderators remain to be fully characterized. Here, we examined the role of encoding variability in the use of recollection- and familiarity-based retrieval processes during lure classifications using experimental and individual differences approaches.

Attention, memory, and individual differences

Relationships between attention and memory may also reveal the role of encoding variability in memory-based lure classifications. We

examined these relationships here by measuring self-reported attention during study and subjective retrieval reports in a mnemonic discrimination task. This allowed us to examine correlations among attention during study, lure classification accuracy, and subjective retrieval experiences evoked by lure features. We have long known that attending to study stimuli promotes better memory (for a review, see Long et al., 2018) and that divided attention during study impairs subsequent memory more for recollection- than familiarity-based retrieval. The latter was shown using process dissociation (e.g., Jacoby, 1998; Jacoby et al., 1993) and remember/know (e.g., Gardiner & Parkin, 1990; McCabe et al., 2011) procedures. Consequently, lures should evoke more recollection-based retrievals when participants had reported attending to studied objects, and this should occur more for participants who better sustain their attention during initial acquisition.

This predicted relationship is supported by studies showing associations between self-reported attention during study and subsequent memory. Higher rates of inattention to study stimuli (i.e., mind wandering) were associated with poorer memory, especially for study tasks that invited semantic processing (Maillet & Rajah, 2013; Thomson et al., 2014). Additionally, in standard verbal learning paradigms, on-task reports during study were associated with more subsequent recollection- than familiarity-based retrieval, whereas off-task reports were associated with no difference between such retrievals (Smallwood et al., 2003, 2007). In contrast, in a natural scene learning paradigm, participants with intermediate mind wandering rates reported the most recollection-based retrievals (Blondé et al., 2020). However, these results are ambiguous because the stimuli afforded participants the opportunity to allocate their attention to extraneous environmental features that were not subsequently tested. Collectively, these findings indicate that self-reported attention is associated with subsequent subjective recollection.

Similar conclusions can be drawn from studies examining associations between self-reported attention and subsequent cued recall, which relies primarily on recollection. One study examined such associations using an A-B, A-D paradigm in which word pairs repeated, appeared alone, or included the same cue with changed responses (Garlitch & Wahlheim, 2020). On-task reports during study were positively associated with target response recall and recollection that responses changed in both within- and between-subjects comparisons. A related study using a paired associate task with self-reports and pupillometry as measures of attention reported similar results (Miller & Unsworth, 2021). Cued recall was higher when participants reported being on– than off-task during study. This improved recall was associated with larger increases in pupil diameter that reflected attentional intensity during encoding (also see Miller & Unsworth, 2019). These findings converge with the aforementioned studies that used subjective retrieval measures in suggesting that sustained attention during encoding supports recollection. Importantly, this collection of studies reveals such associations both within-participants and by leveraging individual differences.

Other studies have used individual differences approaches to identify the roles of retrieval processes in mnemonic discrimination tasks. These approaches entail examining associations between lure classifications and performance on tasks that are sensitive to control processes that govern attention during encoding and strategic retrieval (analogous to recollection). The primary assumption in these studies is that positive between-subject associations between lure rejections and task performance indicates that people who are better at controlling encoding and/ or retrieval are also better at classifying lures. Studies have shown that lure classification accuracy positively associates with free recall performance (e.g., Migo et al., 2014; Toner et al., 2009; Trelle et al., 2017), suggesting that recollection-based retrieval strategies contribute to lure rejection. Other work has shown that lure classification accuracy positively associates with composite scores on tasks of executive functioning and attention that measured working memory, attention and mental flexibility, and verbal fluency (Gellersen et al., 2021, 2023). These associations implicate recollection-based retrieval in lure classifications,

but also suggest a role for controlled attention during encoding. The latter is consistent with findings showing that people who score higher on executive function measures report fewer instances of mind wandering during ongoing tasks (e.g., Kane et al., 2007, 2016; Kane & McVay, 2012; McVay & Kane, 2009). Collectively, these findings converge in suggesting that people who self-report more attentive encoding should be more likely to also report using recollection-based pattern completion when classifying lures.

The present study

We examined encoding variability and retrieval interactions in an object-based mnemonic discrimination task that included thought probes during study and remember/know responses at test. During study, participants considered if objects belonged indoors or outdoors. When probes appeared, participants reported their attentional state as on or off task. *On task* indicated thoughts about object locations, whereas *off task* indicated other thoughts, related or unrelated to the task. We used this method to simplify responding and maximize observations for conditional analyses. This procedure has sensibly captured relative differences in attention and memory in similar tasks (Garlitch & Wahlheim, 2020; Wahlheim et al., 2023), even though it does not capture task-related thoughts that distract from encoding, such as task length and difficulty (e.g., Smallwood et al., 2004; Stawarczyk et al., 2011). We discuss this compromise further in the Discussion section. At test, participants classified studied targets, similar lures, and new foils as old, similar, and new, respectively. After old and similar responses, participants indicated their subjective retrieval experiences by responding "remember" when they recollected studied object details and "familiar" when they felt like they studied that kind of object. We omitted a "guess" response to prevent lax responding and to promote thorough consideration of subjective retrieval states.

We preregistered four hypotheses on the Open Science Framework (OSF). Note that the ordering of hypothesis here and on the OSF differ from the results below. We also report other non-preregistered analyses for completeness. Our first hypothesis is that we will replicate the patterns of within-subjects associations between probe reports and object classifications shown by Wahlheim et al. (2023). We predict that lure rejections (lures called similar), lure false alarms (lures called old), and target hits (targets called old) will be more strongly associated with on– than off-task reports. In contrast, we predict that lure incorrect rejections (lures called new), target incorrect rejections (targets called similar), and target misses (targets called new) will be more strongly associated with off- than on-task reports. The rationale is that on-task reports should reflect attentive encoding that promotes recollection of studied object details.

Our second hypothesis is that we will replicate associations between individual differences in on-task reports and bias-corrected indices shown by Wahlheim et al. (2023). The lure discrimination index corrects for bias by subtracting "similar" classifications for foils from lures [p (lure │ similar) – p(foil │ similar)]. The traditional recognition index corrects for bias by subtracting "old" classifications for foils from targets [p(target │ old) – p(foil │ old)]. If people who are on task more often also encode more effectively, then on-task reports should positively associate with these indices to the extent that memory-based strategies guide classifications. Including retrieval reports at test also allowed us to examine how attentiveness during encoding interacted with recollection- and familiarity-based retrievals. We did not preregister hypotheses about these relationships. However, after preregistering the analyses and before looking at the data, we reasoned that differences in correlation effect sizes would reflect the completeness of attention during encoding. Since full attention promotes recollection-based retrieval (e. g., Jacoby et al., 1993; McCabe et al., 2010) individual differences in on-task reports should more strongly associate with recollection- than familiarity-based object classifications.

Our third hypothesis is based on remember/know rates from two

prior mnemonic discrimination studies (Kim & Yassa, 2013; Szőllősi et al., 2020). We expect to replicate the consistent finding of more recollection than familiar responses for target hits. Based on mixed findings, we did not predict differences in remember and familiar responses for lure rejections and false alarms. However, note that our procedure is more similar to Kim and Yassa (2013). Regardless of the outcome, it is meaningful to examine these response rates. Doing so will provide the literature with additional data points regarding how task designs and participant populations determine the use of memory-based approaches to classifying similar lures.

For the fourth hypothesis, we deviated from our preregistered framing to clarify the predicted relationships among encoding and retrieval states and object classifications. We assumed that on-task reports indicate attentive encoding that promotes recollection-based object classifications. Lure rejections with remember responses will therefore be more strongly associated with on– than off-task reports. In contrast, lure rejections with familiar responses will not be more strongly associated with on-task reports. We also assumed that lure false alarms occur when lure features are unsuccessfully compared with memories for studied objects. Task reports and lure false alarms will therefore not associate with remember or familiar responses. We finally assumed that studied objects are more likely to be recollected after on– than off-task reports. On-task reports should therefore be more strongly associated with target-object recognition for remember but not familiar responses. We did not make a prediction for similar classifications of targets.

Methods

Transparency, openness, and data availability

We report how we determined sample sizes, all data exclusions, all manipulations, and all measures. We preregistered the design and some of the analyses. The present research complied with the Institutional Review Boards at the University of North Carolina at Greensboro (UNCG) and Towson University (UNCG Protocol #IRB-FY22-100; No Protocol # available from Towson). We analyzed the data using R software (R Core Team, 2023) version 4.3, with the packages: *tidyverse*, version 2.0.0 (Wickham et al., 2019); *magrittr*, version 2.0.3 (Bache & Wickham, 2022); *janitor*, version 2.2.0 (Firke, 2023); *lme4*, version 1.1.35.1 (Bates et al., 2015); *car*, version 3.1.2 (Fox & Weisberg, 2019); *emmeans*, version 1.8.9 (Lenth, 2023); *effectsize*, version 0.8.6 (Ben-Shachar et al., 2020); *broom.mixed*, version 0.2.9.5 (Bolker & Robinson, 2024); *moments*, version 0.14.1 (Komsta & Novomestky, 2022); *splithalfr*, version 2.2.2 (Pronk, 2021); and *patchwork*, version 1.1.3 (Pedersen, 2023). The deidentified data, analysis code, and study stimuli are available on the OSF at https://osf.io/fxzb6/.

Participants

Undergraduate students from UNCG and Towson University participated in this study. We preregistered a stopping rule of 300 participants to match the sample in the prior study upon which we based the current study (Wahlheim et al., 2023). We deviated from this rule simply because we ended up with the resources and time to test more participants than anticipated. The following analyses include data from 396 participants, ages 18 – 53 (*M*=19.17, *SD*=2.51) with 198 from UNCG (ages 18 – 28, *M*=19.16, *SD*=1.76), and 198 from Towson (ages 18 – 53, *M*=19.19, *SD*=3.08). According to a sensitivity analysis via G*Power 3.1.9.7 (Faul et al., 2009), we had 80 % power to detect a small effect size (*r* = .14). We tested 407 total participants but did not analyze data from 11 of them because the program failed during testing and did not produce a data file (4 at UNCG; 6 at Towson) or a fire alarm drill ceased the session early (1 at Towson).

Design and materials

The experiment used a within-subjects design with the independent variable Item Type (Similar Lures vs. Target Objects vs. New Foils). The material set comprised 540 pairs of images of everyday objects from the Stark and colleagues database (https://github.com/celstark/MST; Stark et al., 2013). Each pair included two different objects of the same kind (e.g., two rubber ducks). That database provides the normative false alarm rates indicating how often each similar lure was misclassified as a target object. Those rates were used to create bins representing groups of objects varying in the perceptual similarity between targets and lures. From that database, we selected items from lure bins 2–4 to create an intermediate challenge for lure discrimination. We excluded bins 1 and 5, which have the highest and lowest false alarm rates, respectively.

Participants completed two study-test cycles, each with 270 objects. We counterbalanced by creating six groups of 90 objects with comparable normative lure false alarms (*M*=.32, *SD*=.08) and objects from each bin (25 or 26 from Bin 2, 40 from Bin 3, and 25 from Bin 4). We rotated item groups through each item type-by-cycle combination. Groups appeared equally often in each combination across participants, resulting in six unique versions of the experiment.

Fig. 1 displays a procedure schematic. Participants viewed 180 objects in each study phase (360 total) and 270 objects in each test phase (540 total). Each studied object corresponded to a target or lure at test. The tests also included new foils that did not appear during study. Objects appeared in a fixed random order. No more than three of the same item type appeared consecutively. The average serial position was equated across item types. Ten thought probes appeared pseudo-randomly in each study phase (20 total); five appeared for each item type in each phase. The number of objects between probes varied to reduce the predictability of their appearance. Probe intervals were 16, 17, 18, 19, or 20 objects. One probe appeared at each interval for each item type. The average probe interval time was 72 s (*SD*=5.96, range 64–80).

Procedure

We tested participants in groups of 1–4 with an experimenter present. E-prime 3.0 software (Psychology Software Tools, Inc) controlled stimulus presentation. Both study-test cycles used the same procedure with different items. The second cycle started immediately after the first. The full session lasted around 1.5 h.

Each study object appeared for 3 s with a 1 s interstimulus interval (ISI). We chose a longer-than-typical study duration to create optimal thought probe intervals and to avoid ceiling or floor effects in test classifications. We omitted overt judgments to allow more variability in attention during study. But we still encouraged semantic processing with instructions to imagine whether objects belong indoors or outdoors. We also told participants that they would report their attentional states when probes appeared. Participants read the following instructions: "While you are completing this task, you may notice that your ability to focus on the object and imaging it in an indoor or outdoor context waxes and wanes throughout this period. It is normal for people to experience various levels of engagement with a task. We are interested in the extent to which you experience these variations in task engagement. Every now and then, we will ask you to indicate your task engagement during the upcoming study phase. To measure this, we will randomly present a screen that asks you to indicate whether you were on task or off task. If you were thinking about the object in an indoor or outdoor context just before the screen appears, then indicate that you were ON task. In contrast, if your attention was on something else other than imagining the object in this specific context, then indicate that you were OFF task. You will indicate this by clicking on the appropriately labeled button on the screen." Probes appeared after study objects and before ISIs. Participants clicked "On task" and "Off task" buttons on the screen.

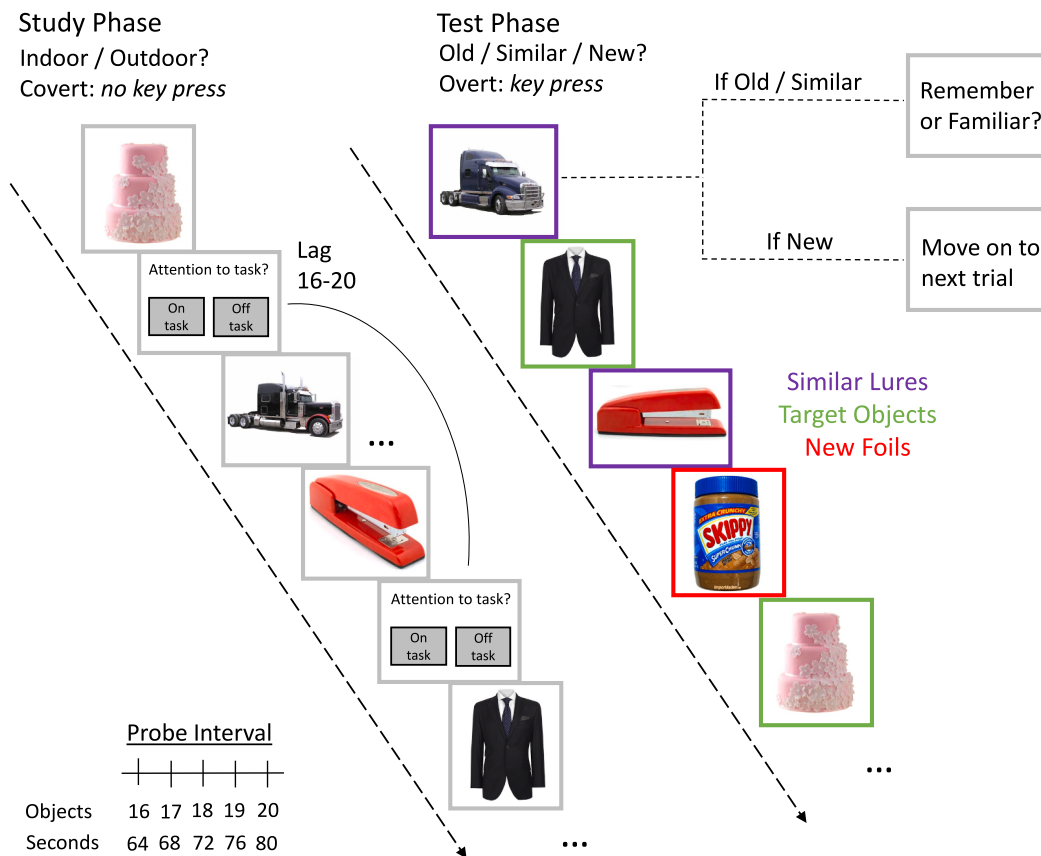The test items included lures, targets, and foils. Lures were

Fig. 1. *Schematic of the Procedure*. During study, participants viewed pictures of objects and covertly considered if objects belonged indoors or outdoors. Thought probes appeared after repeated target and similar lure objects at lags varying from 16 to 20 objects. Probes appeared immediately following objects and comprised "On task" and "Off task" response options. At test, participants viewed pictures of objects and overtly classified them as old, similar, or new with a key press. For items classified as old or similar, participants then made a remember or familiar response to indicate if they could recollect the studied object (remember) or not (familiar). Participants completed two identical study-test cycles with unique object pictures in each.

alternative versions of studied objects. Targets were repetitions of studied objects. Foils were objects that did not appear during the study. Each object appeared initially for 3 s. During that time, participants pressed keys to classify targets as "old" (V), lures as "similar" (B), and foils as "new" (N). When participants responded "new," the program advanced to the next trial. When participants responded "old" or "similar," the object appeared for 3 s longer so that participants could press a key to respond "remember" (R) or "familiar" (F) to indicate their subjective retrieval experience.

Before starting the test phase, participants read the following instructions about how to report their subjective retrieval experiences: "We are also interested in how you judged each object. To recognize an object that you saw before or notice that an object was similar to an earlier object, you can compare it to what you saw before. Sometimes you may clearly remember all the details of an object, other times the details of an object may be fuzzy even when it seems familiar to you. After you indicate that an object was "OLD" or "SIMILAR," we will ask you to describe your memory for the corresponding object you saw earlier. When you clearly remember the details of the earlier object, you should respond that you REMEMBER the earlier object. In contrast, when the earlier object seems familiar, but you cannot remember the specific details, you should respond that the earlier object is FAMILIAR.

When you respond "OLD" by pressing the "V" key, you will be asked to indicate if you REMEMBER the details of the earlier object by pressing the "R" key, or if you are FAMILIAR with the object but do not remember its details from before by pressing the "F" key. When you respond "SIMILAR" by pressing the "B" key, you will also be asked if you REMEMBER the details from the similar object that you saw before or if

the earlier object is only FAMILIAR and you cannot remember its details by pressing the corresponding "R" or "F" key. In both cases, you should respond REMEMBER by pressing the "R" key only if you can consciously recollect seeing the original object on the screen. Otherwise, respond FAMILIAR by pressing the "F" key. You can tell that you consciously recollect an object if you REMEMBER such things as its physical appearance, how it was presented, or perhaps what you thought of when you saw it. However, if you have a strong feeling that the object was previously shown on the screen, but you cannot remember anything about its presentation, respond FAMILIAR. If you respond "NEW" to a test item, you will move onto the next test trial.".

After participants read these instructions, they were shown example images to help them visualize the procedure and to improve their comprehension of when they should respond "remember" and "familiar." After participants finished reading the instructions an experimenter summarized the instructions aloud. Participants were then given the opportunity to discuss the task with an experimenter to resolve any confusion before starting the test phase.

Statistical methods

We examined the effects of interest using standard linear regression and logistic mixed-effects models, depending on the type of outcome variable. The mixed-effects models included random-intercept effects of subject and items. We performed hypothesis tests using the Anova function of the car package (Fox & Weisberg, 2019) and pairwise comparisons controlling for multiple comparisons using the Tukey method from the emmeans function of the emmeans package (Lenth,

2021). Note that the Anova function calculates analysis of variance (type 2) tables for standard linear regression models and Wald's χ^2 for logistic mixed-effect models. We report effects sizes using partial eta squared (η_p^2) and Cohen's d (*d*) estimates for linear regression results and odds ratios (*OR*) for mixed effects model results. We computed bivariate correlations using the cor.test function in base R and report the resulting Pearson r (*r*) effect size estimates. The significance level was α = .05. The code including model specifications and effect size estimation are in an analysis script available on the OSF: https://osf.io/fxzb6/.

Results

Thought probe reports

Fig. 2 displays a characterization of the variation in attention reported on thought probes during the study phases across probe order and participants. The overall patterns closely replicated our previous findings (Wahlheim et al., 2023). In Cycle 1, most participants reported that they were attending as instructed at the outset of the first study phase, but consistently fewer participants reported being on task as the study phase progressed (panel A). The break between cycles reinvigorated attention to studied items at the start of Cycle 2, but again, consistently fewer participants reported being on task across subsequent probes. The overall frequency of on-task report percentages varied across participants (panel B). The vast majority of participants reported following the task instructions more often than not ($\geq$60 % of probes), whereas a relatively small minority reported the opposite. Importantly, the overall variability across participants is ideal for examining within- and between-participant associations between on-task reporting and subsequent object classifications.

Test item classifications

Fig. 3 displays a characterization of object classifications and subjective retrievals at test. The overall classification probabilities (Fig. 3A) indicate that participants followed the task instructions, as classifications were more accurate than not. We did not perform statistical comparisons here because there was no theoretical reason to do so. We present the overall classification probabilities (large dark points) and accompanying between-participant variation (small gray points) to contextualize subsequent analyses of key classification responses.

The goal of our first series of analyses was to characterize the combinations of classifications and subjective retrievals that comprised the bias-corrected indices. This allowed us to examine the extent to which the task evoked specific retrieval processes, and to compare the patterns of those retrievals with prior studies (Kim & Yassa, 2013; Szőllősi et al., 2020). These analyses test for differences in responses rates (Fig. 3B) *before* calculating bias-corrected index scores (Fig. 3C). The first analyses therefore include similar and old classifications with remember and familiar retrievals (Fig. 3B, orange and blue points). We used separate models for similar and old classifications with Item Type and Retrieval Report as predictors. One model compared similar classifications of lures and foils (Fig. 3B, top row, orange points) that eventually comprised the lure discrimination index. Another model compared old classifications of targets and foils (Fig. 3B, bottom row, blue points) that eventually comprised the traditional recognition index. For completeness, we also conducted pairwise comparisons for classifications that did not enter into calculations of bias-corrected indices.

The model for similar classifications of lures and foils indicated a significant effect of Item Type, $F(1, 1564) = 784.24, p < .001, \eta_p^2 = .33$, showing more similar classifications of lures than foils. The model also indicated a significant effect of Retrieval Report, $F(1, 1564) = 9.01, p < .01, \eta_p^2 < .01$, and a significant interaction, $F(1, 1564) = 72.30, p < .001, \eta_p^2 = .04$. Participants made significantly more remember than familiar responses to lures called similar, $t(1564) = 3.85, p < .001, d = 0.27$, consistent with Kim and Yassa (2013) and inconsistent with (Szőllősi et al., 2020). Participants also made significantly more familiar than remember responses to foils called similar, $t(1564) = 8.16, p < .001, d = 0.59$. Finally, a model of targets called similar indicated significantly more familiar than remember responses, $t(790) = 16.26, p < .001, d = 1.16$.

The model for old classifications of targets and foils indicated a significant effect of Item Type, $F(1, 1564) = 2863.10, p < .001, \eta_p^2 = .65$, showing more old classifications of targets than foils, a significant effect of Retrieval Report, $F(1, 1564) = 1948.60, p < .001, \eta_p^2 = .55$, and a significant interaction, $F(1, 1564) = 1926.50, p < .001, \eta_p^2 = .55$. Participants made significantly more remember than familiar responses to targets called old, $t(1564) = 62.25, p < .001, d = 4.94$, replicating earlier findings (Kim & Yassa, 2013; Szőllősi et al., 2020). Participants also made comparable remember and familiar responses when calling foils old, $t(1564) = 0.14, p = .89, d = 0.50$; however, note that floor performance makes this comparison ambiguous to interpret. Finally, a
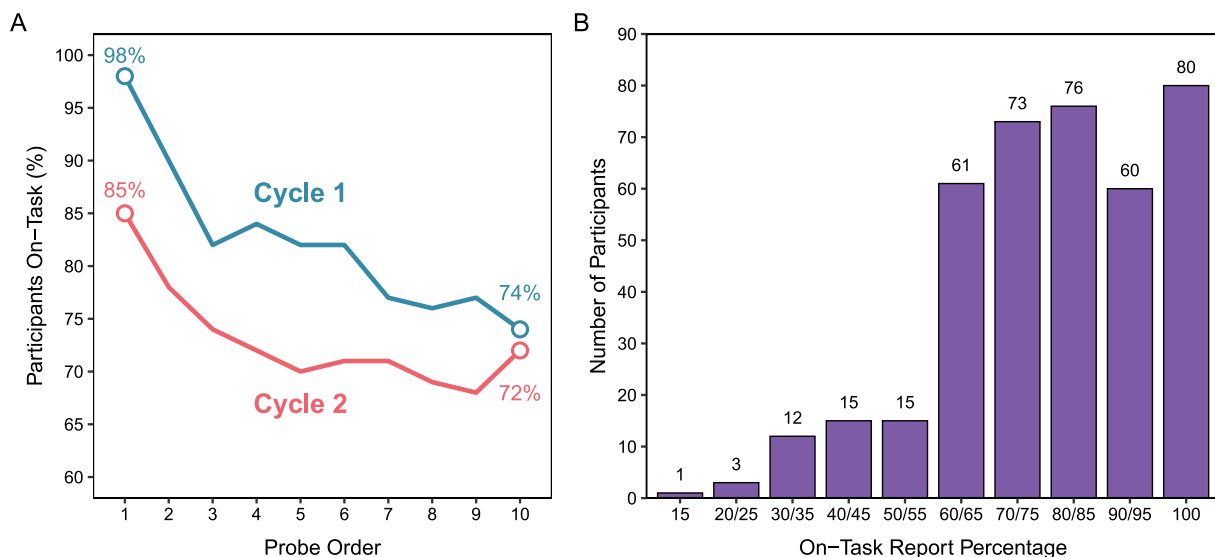

Fig. 2. *On-Task Report Percentages Across Thought Probes and Participants.* (A) The percentages of participants who reported being on task during study conditioned on thought probe order and study-test cycle. (B) The frequency distribution of overall on-task report percentages across participants. The percentages in all bins were multiples of five and the slashes indicate that two values were collapsed into one bin.
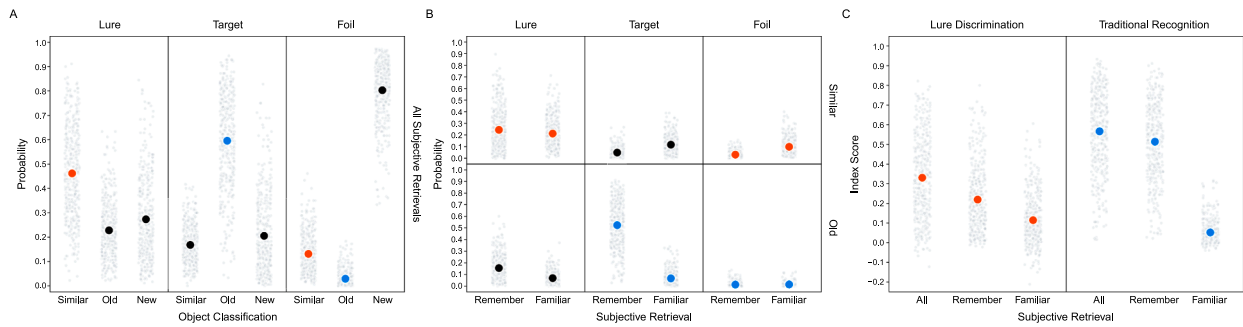
Fig. 3. *Proportions of Object Classifications at Test.* (A) Test classifications for all combinations of object and classification types. (B) Only similar and old test classifications for all item types conditioned on subjective retrieval responses. (C) Bias-corrected index scores for lure discrimination and traditional recognition for all responses and conditioned on subjective retrieval responses. In all panels, the larger colored points are probabilities averaged across participants, whereas the smaller gray points are individual participant probabilities. The orange and blue colorations in panels A and B indicate the cells comprising the lure discrimination and traditional recognition indices, respectively. Those colorations appear again in panel C to show the correspondence with the preceding panels. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

model comparing lures called old indicated that participants reported significantly more remember than familiar responses, $t(790) = 14.83, p < .001, d = 1.05$, replicating earlier findings (Kim & Yassa, 2013; Szőllősi et al., 2020).

These patterns suggest that participants relied more on recollection- than familiarity-based retrieval of studied objects when rejecting lures and endorsing targets. We examined these bases for classifications further by computing lure discrimination and traditional recognition indices (Fig. 3C). We removed response bias by subtracting the probability of foils called similar from lures called similar [p(similar | lure) – p(similar | foil)] and the probability of foils called old from targets called old [p(old | target) – p(old | foil)]. Lure discrimination index scores were significantly higher for remember than familiar responses, $t(774) = 9.71, p < .001, d = 0.70$. Traditional recognition index scores were also significantly higher for remember than familiar responses, $t(774) = 44.75, p < .001, d = 3.21$. These results confirm that lures and targets were more often classified as such using recollection- than familiarity-based retrieval of studied items.

Test Item classifications conditioned on thought probe reports

Fig. 4 shows the within-subjects associations between thought probe reports during study and classifications of lure and target objects at test. We separately modeled differences in each classification response between on– and off-task reports. We did not have complete data for the entire sample because some participants reported being on-task for every probe. However, we included all participants in the analyses by using separate logistic mixed effects models for each object and classification type combination. The models included random-intercept effects of thought probes (within which items were nested) and subjects as well a fixed effect of Probe Report (for another use of this approach, see Miller & Unsworth, 2021). Note that we use this approach for all subsequent analyses that condition outcome probabilities on task reports.

For lures, Fig. 4A shows significantly higher similar responses associated with on– than off-task reports, z ratio = 5.86, $p < .001$, OR=0.59, significantly higher old responses associated with on– than off-task reports, z ratio = 3.98, $p < .001$, OR=0.66, and significantly higher new responses associated with off- than on-task reports, z ratio = 9.78, $p < .001$, OR=2.65. For targets, Fig. 4B shows significantly higher old responses associated with on– than off-task reports, z ratio = 10.06, $p <$
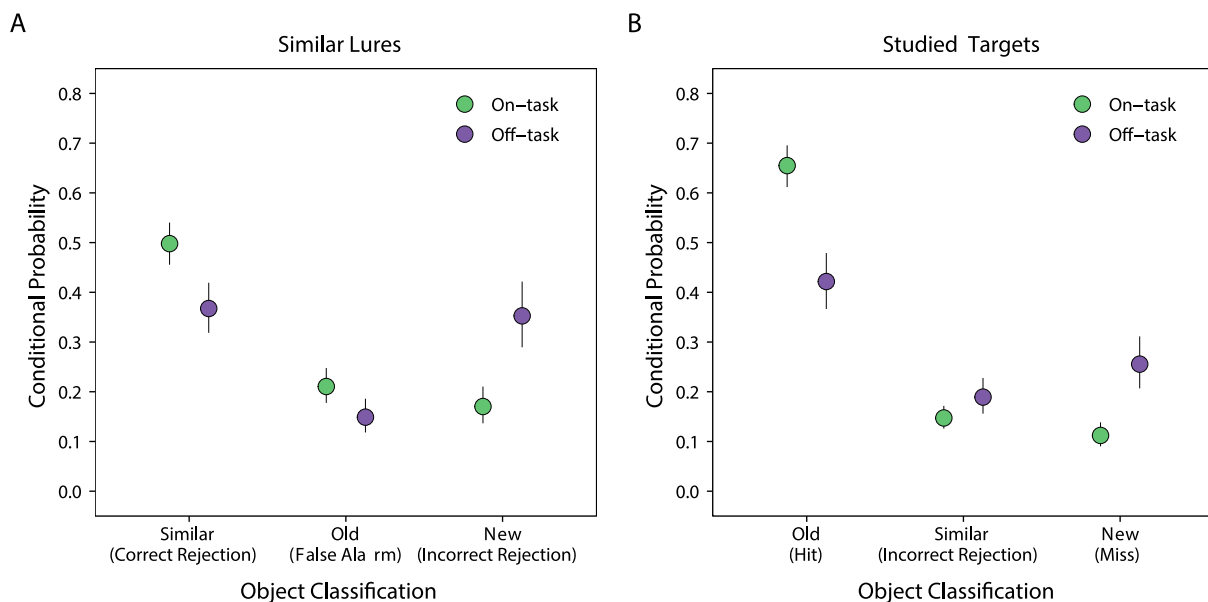

Fig. 4. *Lure and Target Classification Probabilities Conditioned on Thought Probe Reports.* Test response classification probabilities for similar lures (A) and studied targets (B) based on whether participants reported being on task or off task during the study phase. Error bars are 95% confidence intervals.

.001, $OR=0.38$, significantly higher similar responses associated with off- than on-task reports, z ratio $= 2.88$, $p < .01$, $OR=1.35$, and significantly higher new responses associated with off- than on-task reports, z ratio $= 9.31$, $p < .01$, $OR=2.72$. These findings suggest that attention during study promoted memory-based classification strategies for lures and targets.

Test item classifications conditioned on thought probe and subjective retrieval reports

The findings thus far suggest that lure and target classifications reflected memory for studied objects. Also, the retrieval processes supporting these classifications depended on attention during study. Because full attention during study promotes recollection (e.g., Jacoby et al., 1993), we examined whether on-task reports were associated with more recollection- than familiarity-based responses for all test items called similar or old. We used separate logistic mixed effects models to compare probe reports for remember and familiar responses. Fig. 5 shows that attentive study promoted recollection. Remember responses were significantly higher for on– than off-task reports, z ratio $= 13.28$, $p < .001$, $OR=0.40$, whereas familiar responses were significantly higher for off- than on-task reports, z ratio $= 1.98$, $p = .048$, $OR=1.15$.

We then examined associations between attention during study and retrieval processes associated with lure and target classifications to characterize the role of memory-based strategies in object classifications. Fig. 6 shows all combinations of probe and retrieval reports for similar and old classifications of lures and targets. We modeled each classification separately using fixed effects of Probe Report and Retrieval Report. Because there were only two classifications that evoked retrieval reports, the binary outcomes were inverted across models. This led to the same model statistics for similar and old responses within each test item type.

Similar Lures. The model for similar responses to lures (Fig. 6A, left) indicated no significant effect of Probe Report, $\chi^2(1) = 0.33$, $p = .57$, $OR=0.74$, and a significant effect of Retrieval Report, $\chi^2(1) = 84.32$, $p < .001$, $OR=0.23$, showing a lower probability of remember than familiar responses. A significant interaction, $\chi^2(1) = 7.31$, $p < .01$, $OR=1.97$, indicated a higher probability of remember responses for on– than off-task reports, z ratio $= 2.20$, $p = .03$, $OR=0.69$, and no difference in

familiar responses, z ratio $= 1.61$, $p = .11$, $OR=1.36$. These results suggest that attention during study promoted recollection-based lure rejections.

The model for old responses to lures (Fig. 6A, right) indicated no significant effect of Probe Report, $\chi^2(1) = 0.33$, $p = .57$, $OR=1.36$, and a significant effect of Retrieval Report, $\chi^2(1) = 84.32$, $p < .001$, $OR=4.33$, showing a higher probability of remember than familiar responses. A significant interaction, $\chi^2(1) = 7.31$, $p < .01$, $OR=0.51$, indicated a higher probability of remember responses for off- than on-task reports, z ratio $= 2.20$, $p = .03$, $OR=1.45$, and no difference in familiar responses between probe reports, z ratio $= 1.61$, $p = .11$, $OR=0.74$. These results suggest that less attentive encoding promoted recollection of details that were less useful for identifying the different features of test lures.

Studied Targets. The model for old responses to targets (Fig. 6B, left) indicated significant effects of Probe Report, $\chi^2(1) = 4.17$, $p = .04$, $OR=1.32$, and Retrieval Report, $\chi^2(1) = 557.47$, $p < .001$, $OR=19.63$, and no significant interaction, $\chi^2(1) = 0.01$, $p = .93$, $OR=1.02$. The probabilities were higher for remember than familiar responses and for on– than off-task reports. These results suggest that target recognition was more often based on recollection and that attentive encoding promoted both recollection and familiarity-based target recognition.

The model for similar responses to targets (Fig. 6B, right) indicated significant effects of Probe Report, $\chi^2(1) = 4.17$, $p = .04$, $OR=0.76$, and Retrieval Report, $\chi^2(1) = 557.47$, $p < .001$, $OR=.05$, and no significant interaction, $\chi^2(1) = 0.01$, $p = .93$, $OR=.98$. The probabilities were higher for familiar than remember response and for off- than on-task reports. These results are compatible with the conclusion from the patterns of old responses to targets in suggesting that less attentive encoding promoted less detailed retrievals of studied objects.

Individual Differences in Probe Reports and Memory Measures.
The within-subjects associations above and studies showing between-subject associations among executive functioning, memory, and mnemonic discrimination measures (e.g., Gellerson et al., 2021, 2023; Migo et al., 2014; Toner et al., 2009; Trelle et al., 2017) suggest that individual differences in on-task reports should also vary with test responses. Those findings converge in suggesting that people who better focus their attention should be more likely to engage in recollection-based retrieval. We characterized those associations by computing correlations between on-task report percentages and 1) every combination of similar and old classifications for lure and target responses (for all items and only for items given remember or familiar responses), as well as 2) lure discrimination and traditional recognition index scores.

Table 1 displays the descriptive statistics for the probe report and memory measures. We measured internal consistency using the Spearman Brown reliability coefficient. We calculated these correlation coefficients for each measure by computing two scores for each participant using odd and even trials. All measures had reasonable values for internal consistency, except for the bias-corrected indices conditioned on remember and familiar responses. The lower reliability of the index measures likely reflected those measures including two response types for which there were often sparse observations. Many measures were approximately normally distributed. All skewness values were acceptable (<2). Most kurtosis values were acceptable (<4) except for some measures conditioned on remember/familiar responses (see Kline, 1998). The leptokurtic distributions for the latter are not surprising given the narrower response ranges.

We first examined the correlations between on-task reports and lure classifications. These correlations inform our primary interest about the role of attentive study in recollection-based lure classifications. Fig. 7 (top row) shows that a positive association between on-task reports and lure rejections (similar responses) was driven selectively by recollection-based retrievals. On-task reports were significantly associated with lure rejections for all responses (panel A), $r(394) = .33$, $p < .001$, and for only remember responses (panel B), $r(394) = .35$, $p < .001$, but not for only familiar responses (panel C), $r(394) = .01$, $p = .83$. Similarly, Fig. 7 (bottom row) shows that a positive association between on-task reports
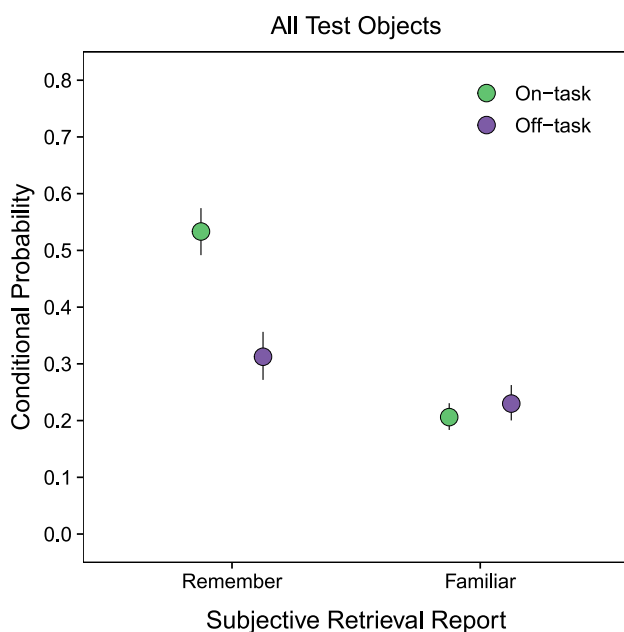
All Test Objects

Fig. 5. *Remember and Familiar Retrieval Probabilities Conditioned on Thought Probe Reports.* Error bars are 95% confidence intervals.
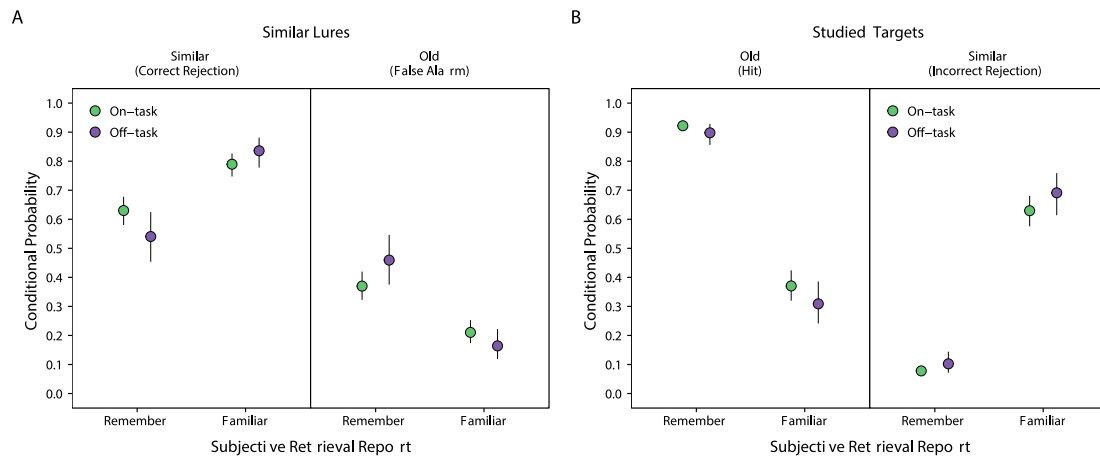
Fig. 6. *Lure and Target Classification Probabilities Conditioned on Thought Probe and Subjective Retrieval Reports*. Test response classification probabilities for (A) similar lures and (B) studied targets. Error bars are 95% confidence intervals and are not visible when they are smaller than point diameters.

Table 1
Descriptive statistics and reliability estimates for probe report and memory measures.

Measure	M	SD	Skew	Kurtosis	Reliability
On-Task	77.53	18.83	−0.72	2.98	.80
Lure Similar (All)	.46	.19	0.10	2.22	.97
Lure Similar (R)	.24	.18	0.90	3.33	.95
Lure Similar (F)	.21	.13	0.94	3.81	.80
Lure Old (All)	.23	.12	0.82	3.63	.94
Lure Old (R)	.16	.10	1.15	4.42	.95
Lure Old (F)	.07	.05	1.76	7.98	.80
Target Old (All)	.60	.19	−0.49	2.66	.97
Target Old (R)	.52	.20	−0.27	2.44	.89
Target Old (F)	.07	.06	1.91	7.60	.72
Target Similar (All)	.17	.08	0.52	2.81	.88
Target Similar (R)	.05	.05	1.76	7.85	.89
Target Similar (F)	.12	.07	0.94	4.07	.72
LDI (All)	.33	.20	0.20	2.24	.95
LDI (R)	.22	.17	0.90	3.29	.31
LDI (F)	.11	.13	0.89	4.06	.45
TRec (All)	.57	.20	−0.51	2.65	.97
TRec (R)	.51	.20	−0.29	2.45	.61
TRec (F)	.05	.06	1.78	6.95	.52

Note. "All" indicates proportions across all items; "R" indicates proportions for only remember responses; "F" indicates proportions for only familiar responses; On-Task, on-task report percentage; Lure Similar, lures called similar; Lure Old, lures called old; Target Old, targets called old; Target Similar, targets called similar; LDI, lure discrimination index score; TRec, traditional recognition index score. Note that because of missing observations, the reliability estimates could not be computed for subsets of participants for LDI (R) (n = 22), LDI (F) (n = 2), TRec (R) (n = 22), and TRec (F) (n = 3).

and lure false alarm (old responses) was also driven selectively by recollection-based retrievals. On-task reports were significantly associated with lure false alarms for all responses (panel D), $r(394) = .13, p = .01$, and for only remember responses (panel E), $r(394) = .14, p < .01$, but not for only familiar responses (panel F), $r(394) = .04, p = .49$. However, comparison of the lure rejections and false alarms associated with remember responses that produced significant correlations (panels B and E) indicated a larger effect size for rejections than false alarms. This difference suggests that recollection of studied objects (viz. pattern completion) was more often a valid basis for lure classifications.

Next, we verified that attentive study was selectively associated with recollection-based retrieval during test item classifications by computing correlations between on-task reports and target classifications. As for lures, Fig. 8 (top row) shows that a positive association between on-task reports and target hits (old responses) was driven selectively by recollection-based retrievals. On-task reports were

significantly associated with target hits for all responses (A), $r(394) = .44, p < .001$, and for only remember responses (B), $r(394) = .46, p < .001$, but not for only familiar responses (C), $r(394) = -.04, p = .41$. In contrast, in the absence of recollection-based retrieval, on-task reports were negatively associated with target classifications that some have considered as akin to lower confidence responses (targets called similar; Loiotile & Courtney, 2015). Fig. 8 (bottom rows) shows no significant association between on-task reports and incorrect target rejections (similar responses) for all responses (D), $r(394) = -.09, p = .08$, and for only remember responses (E), $r(394) = .08, p = .10$; however, there was a significant negative association for only familiar responses (F), $r(394) = -.14, p < .01$. These results suggest that people who were more attentive during study experienced fewer familiarity-based retrievals when classifying target objects.

Finally, for completeness, we verified the association between attentive encoding and recollection-based retrieval by computing correlations between on-task reports and bias-corrected indices. However, note that the indices conditioned on remember/familiar responses were the least reliable of the memory measures (see Table 1). Fig. 9 shows that on-task reports were significantly and positively associated with index scores including all responses for lure discrimination (panel A), $r(394) = .35, p < .001$, and traditional recognition (panel D), $r(394) = .44, p < .001$, replicating earlier findings (Wahlheim et al., 2023). Similar positive associations also obtained for index scores including only remember responses for lure discrimination (panel B), $r(380) = .36, p < .001$, and traditional recognition (panel E), $r(380) = .45$ (panel E), $p < .001$. In contrast, on-task reports were not significantly associated with index scores including only familiar responses for lure discrimination (panel C), $r(392) = .08, p = .09$, and traditional recognition (panel F), $r(392) = .03, p = .61$. The latter two findings converged with the findings above in suggesting that attentive encoding selectively supported recollection of studied items.

Discussion

The present study used an experimental and individual differences approach to examine the roles of attention during study and retrieval quality in mnemonic discrimination. On-task reports indicating attention during study varied across participants and were positively associated with lure rejections and false alarms. These associations were stronger for rejections than false alarms, especially when responses were recollection-based. These correlational results are consistent with the view that attention during study can promote subsequent recollection-based pattern completion that supports lure rejection. Moreover, such retrievals may also lead to false alarms that occur when criterial details are not retrieved and compared with perceptual inputs. The present
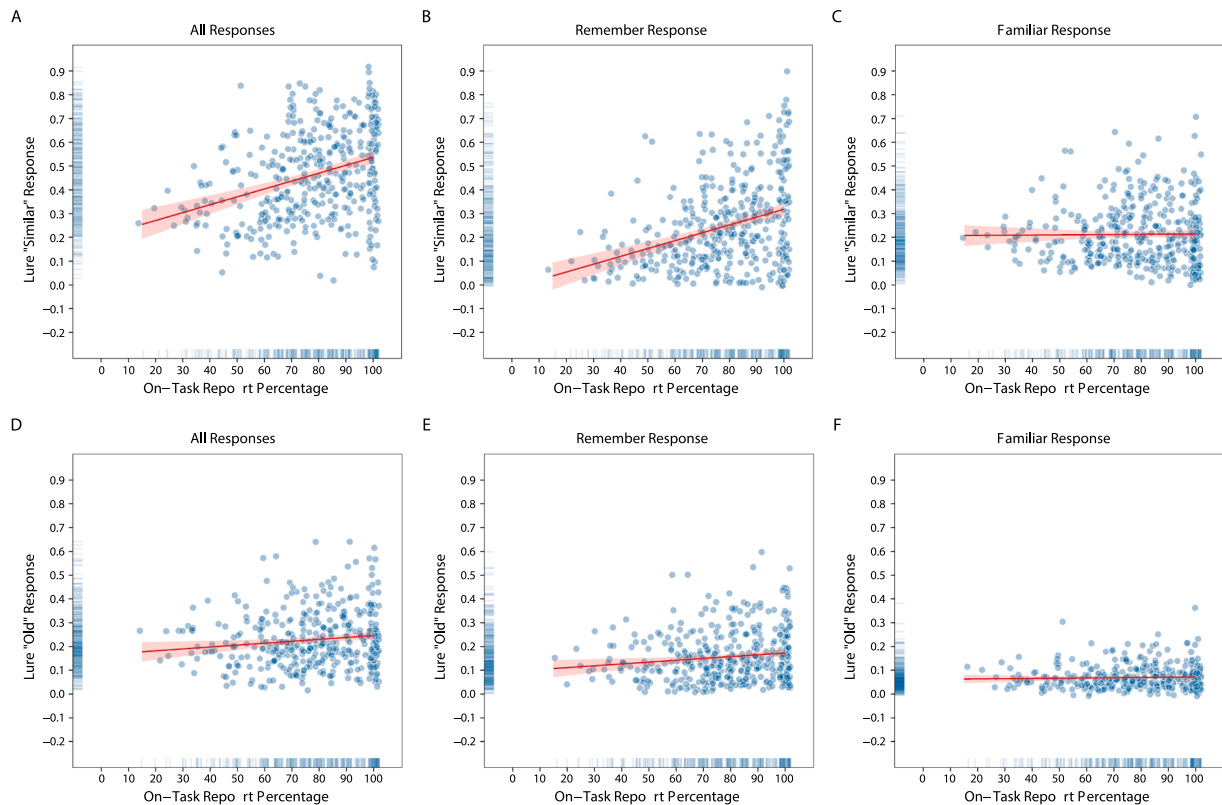
Fig. 7. *Between-Subjects Correlations of On-Task Reports and Lure Classifications.* Scatter plots depicting associations between the percentages of on-task reports during study and lure classifications for "similar" responses (top row) and "old" responses (bottom row). These associations are plotted separately for all test items (A, D), items evoking remember responses (B, E), and items evoking familiar responses. The shaded regions are 95% confidence intervals. The rugs indicate frequency distributions for each measure.

findings emphasize that lure rejection does not purely assay pattern separation in mnemonic discrimination tasks. Pattern completion may also contribute to lure classifications, especially when participants attend effectively during study. Individual differences in controlled attention may therefore determine the balance of pattern separation and pattern completion that occurs when people encode new events that share features with past experiences. We discuss the present findings and their relationship to prior work and theorizing in more detail below.

The present findings are relevant to models of hippocampal computations supporting distinctive encoding to prevent interference (for a review, see Yassa & Stark, 2011). Pattern separation and completion have been examined using recognition tasks in which participants study pictures of category exemplars then discriminate those exemplars from lures that are new exemplars from studied categories (e.g., Koutstaal & Schacter, 1997). Lure rejections were once considered pure reflections of pattern separation, whereas false alarms were considered pure reflections of pattern completion (e.g., Stark et al., 2013). However, lure features can trigger pattern completion that supports rejections, and participants only notice such matches when items were encoded well enough (for a review, see Liu et al., 2016). The present findings contribute to this aspect of the literature by showing that self-reported attention during study is associated with increases in both lure rejections and false alarms, as in Wahlheim et al. (2023). These findings also characterize the associations between self-reported attention during study and subsequent retrieval quality: Attentive encoding promoted recollection- but not familiarity-based lure and target item classifications, which varied across participants.

The associations between probe reports and lure classifications also contribute to the literature on encoding variability and mnemonic discrimination. Those studies have primarily focused on the

contributions of hippocampal subfields DG and CA3 to lure classifications (e.g., Bakker et al., 2008). However, if pattern completion contributes to lure classifications, then encoding variability should create behavioral and neural differences that obscure detection of contributions from pattern separation. This complicates interpretations of pattern separation differences created by experimental manipulations, such as lure-target similarity. For example, attentive study could inflate the contribution of pattern completion from lower-similarity lures that should more naturally evoke pattern separation. Moreover, variation in attentional control across individuals and trials should lead to variation in the balance of pattern completion and separation. Taken with the present findings, these considerations suggest that studies need to account for individual differences in executive functioning associated with variability in controlled attention during encoding to capture the contributions of pattern completion.

Relatedly, as described in the Introduction, multiple approaches suggest that encoding variability leads to a mixture of encoding consequences for lure classifications. These findings are relevant to the debate about whether pattern separation and completion operate independently or compete for hippocampal resources (e.g., Hunsaker & Kenser, 2013; Yassa & Stark, 2011). Studies showing that encoding differences lead to inverse relationships between lure rejections and false alarms suggest that these processes compete for resources (e.g., Huffman & Stark, 2017; Stark et al., 2013; Wahlheim et al., 2022) However, the present results showing that on-task reports were associated with both lure rejections and false alarms contradict this assertion (also see, Wahlheim, 2023). From the shared resource perspective, the present results would be interpreted as showing that attentive encoding was associated with responses that indicate pattern separation (lure rejections) and pattern completion (lure false alarms). Given that the
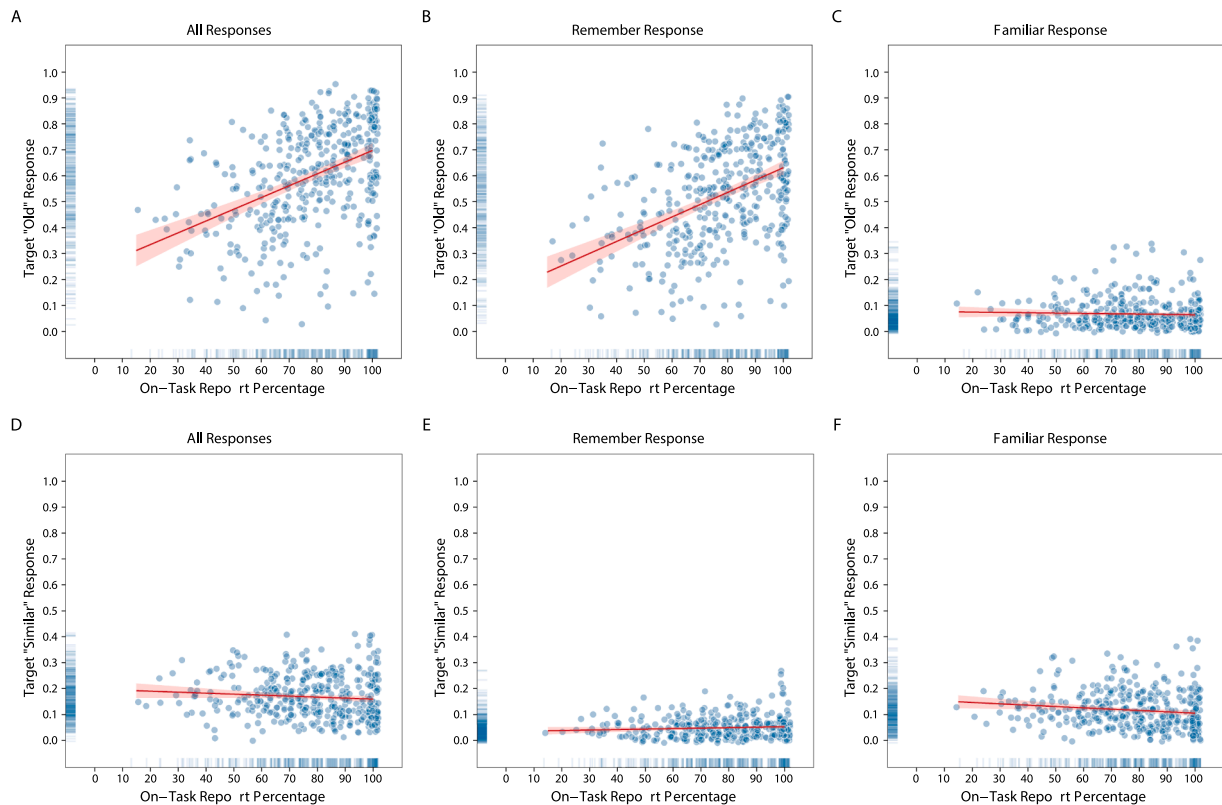
Fig. 8. *Between-Subjects Correlations of On-Task Reports and Target Classifications.* Scatter plots depicting associations between the percentages of on-task reports during study and target classifications for "old" responses (top row) and "similar" responses (bottom row). These associations are plotted separately for all test items (A, D), items evoking remember responses (B, E), and items evoking familiar responses. The shaded regions are 95% confidence intervals. The rugs indicate frequency distributions for each measure.

shared resource perspective was not supported, an alternative possibility is that attentive encoding promoted pattern completion of studied objects, but its consequences for the accuracy of lure classifications depended on the quality of retrieved item features.

We examined associations between encoding variability and pattern completion quality with the same metacognitive measure (i.e., remember/know judgments) used in earlier studies (Kim & Yassa, 2013; Szőllősi et al., 2020). We replicated the finding of lure rejections and false alarms being associated with more recollection- than familiarity-based retrievals (Kim & Yassa, 2013). This was inconsistent with a recent study using a longer test delay and a "guess" option (Szőllősi et al., 2020). That study showed more familiarity- than recollection-based retrievals for lure rejections, and more recollection- than familiarity-based retrievals for lure false alarms. That pattern is the opposite of what one would expect if participants were more likely to reject lures after comparing them with recollected details of studied objects. The longer test delay may have decreased the precision of retrieval and metacognitive resolution (cf. Kelley & Sahakyan, 2003). This could be tested by comparing classifications made immediately and after a delay. Further, we did not include a "guess" option to encourage participants to make old and similar classifications only when they believed an object of the same identity appeared during study. We also wanted to discourage laxed reporting, which is a risk of group testing. This procedural difference was likely trivial for across-study differences because "guess" responses were too infrequent (~7%) to fully account for the discrepant lure rejections in Szőllősi et al. (2020).

Despite the mixed mnemonic bases for lure rejections across studies, lure false alarms have been consistently associated with more recollection- than familiarity-based retrievals. This makes sense because subjective recollection often leads to accurate endorsements of targets (i.e.,

"old" responses). However, recollecting studied object details should also provide a diagnostic basis for rejecting lures by amplifying the salience of discrepant features (also see Gallo, 2004). One account of these recollection-based false alarms is that subjective recollection reports—aggregated across trials—reflect retrievals of criterial and non-criterial features (cf. Parks, 2007; Yonelinas & Jacoby, 1996). Criterial recollection includes features that changed from studied to lure items, thus leading to lure rejections. In contrast, non-criterial recollection includes features shared by studied and lure items that are irrelevant for detecting discrepancies, thus leading to false alarms. The suggestion that recollective content varied across trials here is consistent with the view that recollection- and familiarity-based retrievals vary in strength and content (Wixted & Mickes, 2010). The dichotomous remember/familiar choices here precluded tests of this possibility, but future studies could address this issue using more granular confidence ratings.

More germane to the issues of associations between attention and retrieval quality, on-task reports here were associated more strongly with recollection than familiarity responses. This replicates studies showing that divided attention during study selectively impairs recollection (for a review, see Yonelinas, 2002) and precludes the use of recollection to reject semantic lures (Odegard et al., 2008; Odegard & Lampinen, 2005). Consistent with our proposal that criterial recollection supports lure rejection, on-task reports were also more strongly associated with recollection-based lure rejections. In contrast, off-task reports were more strongly associated with recollection-based false alarms. These associations are consistent with the possibility that attentive encoding led to more precise memories that promoted criterial recollection, whereas inattentive encoding led to less precise memories that promoted non-criterial recollection.

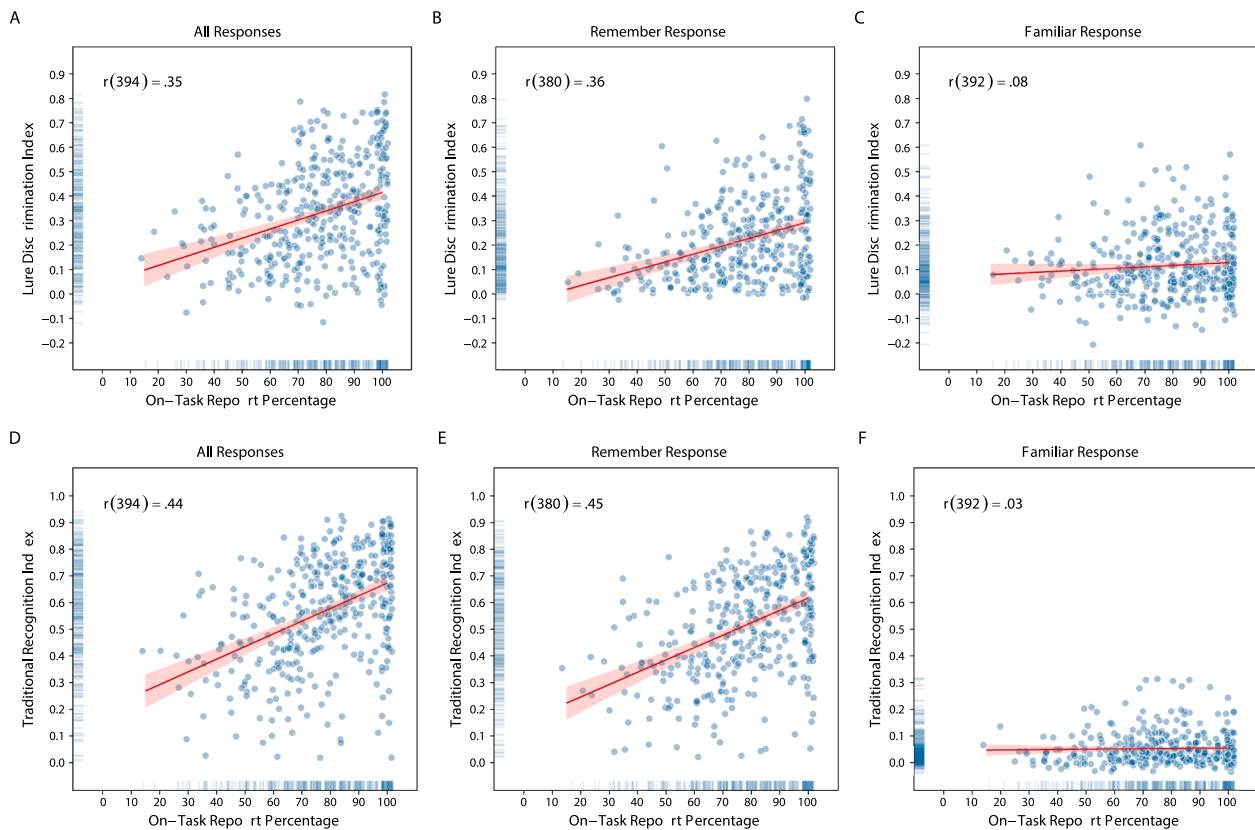Individual differences analyses revealed key support for this

Fig. 9. *Between-Subjects Correlations of On-Task Reports and Bias-Corrected Index Scores.* Scatter plots depicting associations between the percentages of on-task reports during study and index scores for lure discrimination (top row) and traditional recognition (bottom row). These associations are plotted separately for all items (A, D), test items evoking remember responses (B, E), and test items evoking familiar responses (C, F). The shaded regions are 95% confidence intervals. The rugs indicate frequency distributions for each measure.

interpretation. On-task report percentages were more strongly positively associated with recollection-based lure rejections than false alarms, but both correlations were significant. Conversely, on-task reports did not correlate with familiarity-based lure classifications. Together, these results suggest that people who could better focus their attention on objects during study were more often able to use recollection-based pattern completion to detect discrepant lure features. Converging evidence for the association between controlled attention and recollection-based pattern completion was shown by the selective positive correlation between on-task reports and recollection-based target recognition, which strongly reflects pattern completion. Moreover, on-task reports were negatively correlated with familiarity-based "similar" classifications of targets that some have described as low-confidence hits (cf. Loiotile & Courtney, 2015). Those responses should be less common for people who better focus their attention during study. Finally, we replicated the positive correlations between on-task reports and bias-corrected memory indices (Wahlheim et al., 2023). However, the smaller effect size for lure discrimination than traditional recognition suggests that different mechanisms supported lure and target classifications. This contrasts with the suggestion that a common recollection process supports both judgments (e.g., Brainerd et al., 2003; Dobbins et al., 1998; Jones, 2005; Lampinen et al., 2004; Matzen et al., 2011). However, it is compatible with neuroimaging work showing recollection rejection and target recollection are associated with different functional networks of brain regions (Bowman & Dennis, 2016).

Relatedly, the present individual differences in on-task reports and memory outcomes have implications for the neural mechanisms supporting lure classifications. Although mnemonic discrimination tasks are used to examine hippocampal processes, mnemonic discrimination also

involves cortical regions (for a review, see Amer & Davachi, 2023). Of primary relevance here, the literature on attention control and interference has shown interactions among frontoparietal regions when conflict occurs among stimuli or memories (e.g., Amer et al., 2016; Badre & Wagner, 2007). Also, working memory tasks that require control to resolve interference activate prefrontal regions (e.g., Jonides et al., 1998, 2000). Taken with the view that executive function recruits attentional control mechanisms supported by prefrontal regions (Engle & Kane, 2004; McCabe et al., 2010), this suggests that individual differences in on-task reports and lure classifications reflect more than variations in hippocampal pattern separation and completion. Instead, the correlations that we report here suggest that people with more control over their attention may also more effectively recruit networks of brain regions that serve to distinguish retrieved memories from similar perceptions. This possibility is compatible with the finding that a pattern of intrinsic functional connectivity in the default mode network—which includes prefrontal, hippocampal, and other regions—predicts individual differences in mnemonic discrimination operationalized as the lure discrimination index (Wahlheim et al., 2022).

Finally, the associations revealed by individual differences here are also consistent with studies of individual differences in mnemonic discrimination. Studies showed that people who performed better on executive function and strategic retrieval measures also better discriminated similar lures from studied items (e.g., Migo et al., 2014; Gellersen et al., 2021). The present study showed that people who reported more attentive encoding also rejected more lures using controlled, recollection-based retrieval. Whereas prior studies could only infer differences in such encoding/retrieval interactions from control measures,

the present study measured that relationship directly using subjective reports of controlled attention during study and retrieval quality at test. Taken with the broader cognitive and neuroscience literatures on attention/memory interactions and mnemonic discrimination, these findings suggest that people who can better focus attention during study can better recruit high-quality pattern completion, which may reflect individual differences in prefrontal functioning and cortical-hippocampal connectivity.

The present study had several limitations. First, as mentioned in the Introduction, the thought probes included a binary response configuration. We therefore could not capture times when participants thought about the task but were not "on task" per se. Research investigating this issue in attention tasks has shown that including an option to report such task-related thoughts reduced response rates more for on– than off-task reports (Kane et al., 2021; Robison et al., 2019). Therefore, if one assumes that task-related thoughts distract from encoding and impair memory, then the clear associations between on-task reports and memory outcomes observed here may be more conservative than what would obtain using thought probes with three or more response options. If anything, this limitation bolsters the robustness of the present findings. Second, although the distributions and reliabilities of memory measures were mostly suitable for individual differences analyses, there were non-normal distributions and unacceptable reliability for a minority of measures, possibly reflecting sparser observations. This suggests the need for cautious interpretation of associations involving those measures and future replication attempts. Third, although the present findings suggest a role for controlled attention during study in memory-based classifications, we did not include converging measures of executive functions and sustained attention. Future studies could further test our interpretations of the present associations, especially those pertaining to the role of executive functioning and strategic retrieval differences in mnemonic discrimination, by including three or more tasks per construct to provide a latent variable perspective on these relationships.

Conclusion

The present experiment used an experimental and individual differences approach to characterize the associations among self-reported attentional states, retrieval quality, and mnemonic discrimination. The present findings provide indirect evidence consistent with the suggestion that recollection-based pattern completion contributes to lure rejection. Taken with the available literature, these findings further contradict the suggestion that lure rejections are a pure behavioral assay of pattern separation. The associations between self-reported attention and retrieval here suggested that memories of studied items cued by lure features were more likely to include criterial features when participants attended to stimuli effectively during study. The present findings also contribute to the literature on individual differences in mind wandering and memory by showing another instance when attentive study is associated with recollection-based retrieval. Our understanding of the role of controlled attention and retrieval in mnemonic discrimination would benefit from future individual differences studies that combine variants of the present paradigm including thought probes with measures of executive function and strategic retrieval.

CRediT authorship contribution statement

Christopher N. Wahlheim: Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation, Resources, Data curation, Writing – original draft, Writing – review & editing, Visualization, Supervision, Project administration. **Sydney M. Garlitch:** Writing – review & editing, Visualization, Validation, Supervision, Software, Project administration, Methodology, Conceptualization. **Rawan M. Mohamed:** Supervision, Project administration. **Blaire J. Weidler:** Writing – review & editing, Supervision, Project administration,

Investigation, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

I have shared the link to my data.

References

Ally, B. A., Hussey, E. P., Ko, P. C., & Molitor, R. J. (2013). Pattern separation and pattern completion in Alzheimer's disease: Evidence of rapid forgetting in amnestic mild cognitive impairment: Pattern Separation in AMCI and AD. *Hippocampus, 23*(12), 1246–1258. https://doi.org/10.1002/hipo.22162

Amer, T., Campbell, K. L., & Hasher, L. (2016). Cognitive control as a double-edged sword. *Trends in Cognitive Sciences, 20*(12), 905–915. https://doi.org/10.1016/j.tics.2016.10.002

Amer, T., & Davachi, L. (2023). Extra-hippocampal contributions to pattern separation. *eLife, 12*, e82250.

Bache, S., & Wickham, H. (2022). *magrittr: A Forward-Pipe Operator for R* (2.0.3) [Computer software]. https://CRAN.R-project.org/package=magrittr.

Badre, D., & Wagner, A. D. (2007). Left ventrolateral prefrontal cortex and the cognitive control of memory. *Neuropsychologia, 45*(13), 2883–2901. https://doi.org/10.1016/j.neuropsychologia.2007.06.015

Bakker, A., Kirwan, C. B., Miller, M., & Stark, C. E. L. (2008). Pattern separation in the human hippocampal CA3 and dentate gyrus. *Science, 319*(5870), 1640–1642. https://doi.org/10.1126/science.1152882

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software, 67*(1). https://doi.org/10.18637/jss.v067.i01

Ben-Shachar, M., Lüdecke, D., & Makowski, D. (2020). effectsize: Estimation of effect size indices and standardized parameters. *Journal of Open Source Software, 5*(56), 2815. https://doi.org/10.21105/joss.02815

Bjornn, D. K., Van, J., & Kirwan, C. B. (2022). The contributions of eye gaze fixations and target-lure similarity to behavioral and fMRI indices of pattern separation and pattern completion. *Cognitive Neuroscience, 13*(3–4), 171–181. https://doi.org/10.1080/17588928.2022.2060200

Blondé, P., Makowski, D., Sperduti, M., & Piolino, P. (2020). In Medio Stat Virtus: Intermediate levels of mind wandering improve episodic memory encoding in a virtual environment. *Psychological Research*. https://doi.org/10.1007/s00426-020-01358-5

Bolker, B., & Robinson, D. (2024). *broom.mixed: Tidying methods for mixed models* (0.2.9.5) [Computer software]. https://CRAN.R-project.org/package=broom.mixed.

Bowman, C. R., & Dennis, N. A. (2016). The neural basis of recollection rejection: Increases in hippocampal–prefrontal connectivity in the absence of a shared recall-to-reject and target recollection network. *Journal of Cognitive Neuroscience, 28*(8), 1194–1209. https://doi.org/10.1162/jocn_a_00961

Brainerd, C. J., Reyna, V. F., Wright, R., & Mojardin, A. H. (2003). Recollection rejection: False-memory editing in children and adults. *Psychological Review, 110*(4), 762–784. https://doi.org/10.1037/0033-295X.110.4.762

Davidson, P. S. R., Vidjen, P., Trincao-Batra, S., & Collin, C. A. (2019). Older adults' lure discrimination difficulties on the mnemonic similarity task are significantly correlated with their visual perception. *The Journals of Gerontology: Series B, 74*(8), 1298–1307. https://doi.org/10.1093/geronb/gby130

Dobbins, I. G., Kroll, N. E. A., Yonelinas, A. P., & Liu, Q. (1998). Distinctiveness in recognition and free recall: The role of recollection in the rejection of the familiar. *Journal of Memory and Language, 38*(4), 381–400. https://doi.org/10.1006/jmla.1997.2554

Engle, R. W., & Kane, M. J. (2004). Executive attention, working memory capacity, and a two-factor theory of cognitive control. In *Psychology of Learning and Motivation* (Vol. 44, pp. 145–199). Elsevier. https://doi.org/10.1016/S0079-7421(03)44005-X.

Firke, S. (2023). *janitor: Simple tools for examining and cleaning dirty data* (2.2.0) [Computer software]. https://CRAN.R-project.org/package=janitor.

Fox, J., & Weisberg, S. (2019). *An {R} Companion to Applied Regression, Third Edition.* Sage. https://socialsciences.mcmaster.ca/jfox/Books/Companion/.

Gallo, D. A. (2004). Using recall to reduce false recognition: Diagnostic and disqualifying monitoring. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 30*(1), 120–128. https://doi.org/10.1037/0278-7393.30.1.120

Gardiner, J. M., & Parkin, A. J. (1990). Attention and recollective experience in recognition memory. *Memory & Cognition, 18*(6), 579–583. https://doi.org/10.3758/BF03197100

Garlitch, S. M., & Wahlheim, C. N. (2020). The role of attentional fluctuation during study in recollecting episodic changes at test. *Memory & Cognition, 48*(5), 800–814. https://doi.org/10.3758/s13421-020-01018-4

Gellersen, H. M., Trelle, A. N., Farrar, B. G., Coughlan, G., Korkki, S. M., Henson, R. N., & Simons, J. S. (2023). Medial temporal lobe structure, mnemonic and perceptual discrimination in healthy older adults and those at risk for mild cognitive

impairment. *Neurobiology of Aging, 122*, 88–106. https://doi.org/10.1016/j.neurobiolaging.2022.11.004

Gellersen, H. M., Trelle, A. N., Henson, R. N., & Simons, J. S. (2021). Executive function and high ambiguity perceptual discrimination contribute to individual differences in mnemonic discrimination in older adults. *Cognition, 209*, Article 104556. https://doi.org/10.1016/j.cognition.2020.104556

Holden, H. M., & Gilbert, P. E. (2012). Less efficient pattern separation may contribute to age-related spatial memory deficits. *Frontiers in Aging Neuroscience, 4*. https://doi.org/10.3389/fnagi.2012.00009

Huffman, D. J., & Stark, C. E. L. (2017). Age-related impairment on a forced-choice version of the Mnemonic Similarity Task. *Behavioral Neuroscience, 131*(1), 55–67. https://doi.org/10.1037/bne0000180

Hunsaker, M. R., & Kesner, R. P. (2013). The operation of pattern separation and pattern completion processes associated with different attributes or domains of memory. *Neuroscience & Biobehavioral Reviews, 37*(1), 36–58. https://doi.org/10.1016/j.neubiorev.2012.09.014

Jacoby, L. L. (1991). A process dissociation framework: Separating automatic from intentional uses of memory. *Journal of Memory and Language, 30*(5), 513–541. https://doi.org/10.1016/0749-596X(91)90025-F

Jacoby, L. L. (1998). Invariance in automatic influences of memory: Toward a user's guide for the process-dissociation procedure. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 24*(1), 3–26.

Jacoby, L. L., Toth, J. P., & Yonelinas, A. P. (1993). Separating conscious and unconscious influences of memory: Measuring recollection. *Journal of Experimental Psychology: General, 122*(2), 139–154. https://doi.org/10.1037/0096-3445.122.2.139

Jones, T. (2005). Study repetition and the rejection of conjunction lures. *Memory, 13*(5), 499–515. https://doi.org/10.1080/09658210444000197

Jonides, J., Marshuetz, C., Smith, E. E., Reuter-Lorenz, P. A., Koeppe, R. A., & Hartley, A. (2000). Age differences in behavior and PET activation reveal differences in interference resolution in verbal working memory. *Journal of Cognitive Neuroscience, 12*(1), 188–196. https://doi.org/10.1162/089892900561823

Jonides, J., Smith, E. E., Marshuetz, C., Koeppe, R. A., & Reuter-Lorenz, P. A. (1998). Inhibition in verbal working memory revealed by brain activation. *Proceedings of the National Academy of Sciences, 95*(14), 8410–8413. https://doi.org/10.1073/pnas.95.14.8410

Kane, M. J., Brown, L. H., McVay, J. C., Silvia, P. J., Myin-Germeys, I., & Kwapil, T. R. (2007). For whom the mind wanders, and when: An experience-sampling study of working memory and executive control in daily life. *Psychological Science, 18*(7), 614–621. https://doi.org/10.1111/j.1467-9280.2007.01948.x

Kane, M. J., & McVay, J. C. (2012). What Mind Wandering Reveals About Executive-Control Abilities and Failures. *Current Directions in Psychological Science, 21*(5), 348–354. https://doi.org/10.1177/0963721412454875

Kane, M. J., Meier, M. E., Smeekens, B. A., Gross, G. M., Chun, C. A., Silvia, P. J., & Kwapil, T. R. (2016). Individual differences in the executive control of attention, memory, and thought, and their associations with schizotypy. *Journal of Experimental Psychology: General, 145*(8), 1017–1048. https://doi.org/10.1037/xge0000184

Kane, M. J., Smeekens, B. A., Meier, M. E., Welhaf, M. S., & Phillips, N. E. (2021). Testing the construct validity of competing measurement approaches to probed mind-wandering reports. *Behavior Research Methods, 53*(6), 2372–2411. https://doi.org/10.3758/s13428-021-01557-x

Kelley, C. M., & Sahakyan, L. (2003). Memory, monitoring, and control in the attainment of memory accuracy. *Journal of Memory and Language, 48*(4), 704–721. https://doi.org/10.1016/S0749-596X(02)00504-1

Kirwan, C. B., Hartshorn, A., Stark, S. M., Goodrich-Hunsaker, N. J., Hopkins, R. O., & Stark, C. E. L. (2012). Pattern separation deficits following damage to the hippocampus. *Neuropsychologia, 50*(10), 2408–2414. https://doi.org/10.1016/j.neuropsychologia.2012.06.011

Komsta, L., & Novomestky, F. (2022). *moments: Moments, cumulants, skewness, kurtosis and related tests* (0.14.1) [Computer software]. https://CRAN.R-project.org/package=moments.

Koutstaal, W., & Schacter, D. L. (1997). Gist-based false recognition of pictures in older and younger adults. *Journal of Memory and Language, 37*(4), 555–583. https://doi.org/10.1006/jmla.1997.2529

Lampinen, J. M., Odegard, T. N., & Neuschatz, J. S. (2004). Robust recollection rejection in the memory conjunction paradigm. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 30*(2), 332–342. https://doi.org/10.1037/0278-7393.30.2.332

Lenth, R. V. (2023). *emmeans: Estimated marginal means, aka least-squares means* (1.8.9) [Computer software]. https://CRAN.R-project.org/package=emmeans.

Liu, K. Y., Gould, R. L., Coulson, M. C., Ward, E. V., & Howard, R. J. (2016). Tests of pattern separation and pattern completion in humans: A systematic review. *Hippocampus, 26*(6), 705–717. https://doi.org/10.1002/hipo.22561

Loiotile, R. E., & Courtney, S. M. (2015). A signal detection theory analysis of behavioral pattern separation paradigms. *Learning & Memory, 22*(8), 364–369. https://doi.org/10.1101/lm.038141.115

Long, N. M., Kuhl, B. A., & Chun, M. M. (2018). Memory and attention. In J. T. Wixted (Ed.), *Stevens' Handbook of Experimental Psychology and Cognitive Neuroscience* (1st ed., pp. 1–37). Wiley. 10.1002/9781119170174.epcn109.

Maillet, D., & Rajah, M. N. (2013). Age-related changes in frequency of mind-wandering and task-related interferences during memory encoding and their impact on retrieval. *Memory, 21*(7), 818–831. https://doi.org/10.1080/09658211.2012.761714

Matzen, L. E., Taylor, E. G., & Benjamin, A. S. (2011). Contributions of familiarity and recollection rejection to recognition: Evidence from the time course of false

recognition for semantic and conjunction lures. *Memory, 19*(1), 1–16. https://doi.org/10.1080/09658211.2010.530271

McCabe, D. P., Roediger, H. L., & Karpicke, J. D. (2011). Automatic processing influences free recall: Converging evidence from the process dissociation procedure and remember-know judgments. *Memory & Cognition, 39*(3), 389–402. https://doi.org/10.3758/s13421-010-0040-5

McCabe, D. P., Roediger, H. L., McDaniel, M. A., Balota, D. A., & Hambrick, D. Z. (2010). The relationship between working memory capacity and executive functioning: Evidence for a common executive attention construct. *Neuropsychology, 24*(2), 222–243. https://doi.org/10.1037/a0017619

McVay, J. C., & Kane, M. J. (2009). Conducting the train of thought: Working memory capacity, goal neglect, and mind wandering in an executive-control task. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 35*(1), 196–204. https://doi.org/10.1037/a0014104

Migo, E. M., Quamme, J. R., Holmes, S., Bendell, A., Norman, K. A., Mayes, A. R., & Montaldi, D. (2014). Individual differences in forced-choice recognition memory: Partitioning contributions of recollection and familiarity. *Quarterly Journal of Experimental Psychology, 67*(11), 2189–2206. https://doi.org/10.1080/17470218.2014.910240

Miller, A. L., & Unsworth, N. (2019). Variation in attention at encoding: Insights from pupillometry and eye gaze fixations. *Journal of Experimental Psychology: Learning, Memory, and Cognition.*. https://doi.org/10.1037/xlm0000797

Miller, A. L., & Unsworth, N. (2021). Attending to encode: The role of consistency and intensity of attention in learning ability. *Journal of Memory and Language, 121*, Article 104276. https://doi.org/10.1016/j.jml.2021.104276

Molitor, R. J., Ko, P. C., Hussey, E. P., & Ally, B. A. (2014). Memory-related eye movements challenge behavioral measures of pattern completion and pattern separation: Eye Movements and Pattern Completion. *Hippocampus, 24*(6), 666–672. https://doi.org/10.1002/hipo.22256

Nakashiba, T., Cushman, J. D., Pelkey, K. A., Renaudineau, S., Buhl, D. L., McHugh, T. J., Barrera, V. R., Chittajallu, R., Iwamoto, K. S., McBain, C. J., Fanselow, M. S., & Tonegawa, S. (2012). Young dentate granule cells mediate pattern separation, whereas old granule cells facilitate pattern completion. *Cell, 149*(1), 188–201. https://doi.org/10.1016/j.cell.2012.01.046

Ngo, C. T., Michelmann, S., Olson, I. R., & Newcombe, N. S. (2021). Pattern separation and pattern completion: Behaviorally separable processes? *Memory & Cognition, 49*(1), 193–205. https://doi.org/10.3758/s13421-020-01072-y

Odegard, T. N., Koen, J. D., & Gama, J. M. (2008). Process demands of rejection mechanisms of recognition memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 34*(5), 1296–1304. https://doi.org/10.1037/a0013034

Odegard, T. N., & Lampinen, J. M. (2005). Recollection rejection: Gist cuing of verbatim memory. *Memory & Cognition, 33*(8), 1422–1430. https://doi.org/10.3758/BF03193375

Parks, C. M. (2007). The role of noncriterial recollection in estimating recollection and familiarity. *Journal of Memory and Language, 57*(1), 81–100. https://doi.org/10.1016/j.jml.2007.03.003

Pedersen, T. L. (2023). *patchwork: The composer of plots* (1.1.3) [Computer software]. https://CRAN.R-project.org/package=patchwork.

Pronk, T. (2021). *splithalfr: Estimates split-half reliabilities for scoring algorithms of cognitive tasks and questionnaires* (2.2.2) [Computer software]. 10.5281/zenodo.7777894.

R Core Team. (2023). *R: A language and environment for statistical computing* [Computer software]. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/.

Reagh, Z. M., & Yassa, M. A. (2014). Repetition strengthens target recognition but impairs similar lure discrimination: Evidence for trace competition. *Learning & Memory, 21*(7), 342–346. https://doi.org/10.1101/lm.034546.114

Robison, M. K., Miller, A. L., & Unsworth, N. (2019). Examining the effects of probe frequency, response options, and framing within the thought-probe method. *Behavior Research Methods, 51*(1), 398–408. https://doi.org/10.3758/s13428-019-01212-6

Rollins, L., Khuu, A., & Lodi, N. (2019). Encoding variability accounts for false recognition of noncorresponding lures on the forced-choice Mnemonic Similarity Task. *Learning & Memory, 26*(8), 280–283. https://doi.org/10.1101/lm.049593.119

Rolls, E. T. (2013). The mechanisms for pattern completion and pattern separation in the hippocampus. *Frontiers in Systems Neuroscience, 7*. https://doi.org/10.3389/fnsys.2013.00074

Rolls, E. T. (2016). Pattern separation, completion, and categorisation in the hippocampus and neocortex. *Neurobiology of Learning and Memory, 129*, 4–28. https://doi.org/10.1016/j.nlm.2015.07.008

Smallwood, J. M., Baracaia, S. F., Lowe, M., & Obonsawin, M. (2003). Task unrelated thought whilst encoding information. *Consciousness and Cognition, 12*(3), 452–484. https://doi.org/10.1016/S1053-8100(03)00018-7

Smallwood, J. M., Davies, J. B., Heim, D., Finnigan, F., Sudberry, M., O'Connor, R., & Obonsawin, M. (2004). Subjective experience and the attentional lapse: Task engagement and disengagement during sustained attention. *Consciousness and Cognition, 13*(4), 657–690. https://doi.org/10.1016/j.concog.2004.06.003

Smallwood, J. M., McSpadden, M., & Schooler, J. W. (2007). The lights are on but no one's home: Meta-awareness and the decoupling of attention when the mind wanders. *Psychonomic Bulletin & Review, 14*(3), 527–533. https://doi.org/10.3758/BF03194102

Stark, S. M., Kirwan, C. B., & Stark, C. E. L. (2019). Mnemonic smilarity task: A tool for assessing hippocampal integrity. *Trends in Cognitive Sciences, 23*(11), 938–951. https://doi.org/10.1016/j.tics.2019.08.003

Stark, S. M., Yassa, M. A., Lacy, J. W., & Stark, C. E. L. (2013). A task to assess behavioral pattern separation (BPS) in humans: Data from healthy aging and mild cognitive impairment. *Neuropsychologia, 51*(12), 2442–2449. https://doi.org/10.1016/j.neuropsychologia.2012.12.014

Stawarczyk, D., Majerus, S., Maj, M., Van Der Linden, M., & D'Argembeau, A. (2011). Mind-wandering: Phenomenology and function as assessed with a novel experience sampling method. *Acta Psychologica, 136*(3), 370–381. https://doi.org/10.1016/j.actpsy.2011.01.002

Thomson, D. R., Smilek, D., & Besner, D. (2014). On the asymmetric effects of mind-wandering on levels of processing at encoding and retrieval. *Psychonomic Bulletin & Review, 21*(3), 728–733. https://doi.org/10.3758/s13423-013-0526-9

Toner, C. K., Pirogovsky, E., Kirwan, C. B., & Gilbert, P. E. (2009). Visual object pattern separation deficits in nondemented older adults. *Learning & Memory, 16*(5), 338–342. https://doi.org/10.1101/lm.1315109

Trelle, A. N., Henson, R. N., Green, D. A. E., & Simons, J. S. (2017). Declines in representational quality and strategic retrieval processes contribute to age-related increases in false recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 43*(12), 1883–1897. https://doi.org/10.1037/xlm0000412

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

Wahlheim, C. N., Christensen, A. P., Reagh, Z. M., & Cassidy, B. S. (2022). Intrinsic functional connectivity in the default mode network predicts mnemonic discrimination: A connectome-based modeling approach. *Hippocampus, 32*(1), 21–37. https://doi.org/10.1002/hipo.23393

Wahlheim, C. N., Garlitch, S. M., Mohamed, R. M., & Weidler, B. J. (2023). Self-reported encoding quality promotes lure rejections and false alarms. *Learning & Memory, 30,* 96–100.

Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L., François, R., Grolemund, G., Hayes, A., Henry, L., Hester, J., Kuhn, M., Pedersen, T., Miller, E., Bache, S., Müller, K., Ooms, J., Robinson, D., Seidel, D., Spinu, V., & Yutani, H. (2019). Welcome to the tidyverse. *Journal of Open Source Software, 4*(43), 1686. https://doi.org/10.21105/joss.01686

Wixted, J. T., & Mickes, L. (2010). A continuous dual-process model of remember/know judgments. *Psychological Review, 117*(4), 1025–1054. https://doi.org/10.1037/a0020874

Yassa, M. A., & Stark, C. E. L. (2011). Pattern separation in the hippocampus. *Trends in Neurosciences, 34*(10), 515–525. https://doi.org/10.1016/j.tins.2011.06.006

Yonelinas, A. P. (2002). The nature of recollection and familiarity: A review of 30 years of research. *Journal of Memory and Language, 46*(3), 441–517. https://doi.org/10.1006/jmla.2002.2864

Yonelinas, A. P., & Jacoby, L. L. (1996). Noncriterial recollection: Familiarity as automatic, irrelevant recollection. *Consciousness and Cognition, 5*(1–2), 131–141. https://doi.org/10.1006/ccog.1996.0008

Zhang, M., & Hupbach, A. (2020). Repeated encoding fosters retention of perceptual detail in visual recognition memory. *Learning & Memory, 27*(11), 457–461. https://doi.org/10.1101/lm.052209.120

Zhang, M., & Hupbach, A. (2023). The effects of variable encoding contexts on item and source recognition. *Memory & Cognition, 51*(2), 391–403. https://doi.org/10.3758/s13421-022-01353-8