



The effects of mnemonic variability and spacing on memory over multiple timescales

Emily T. Cowan^{a,1} , Yiwen Zhang^b , Benjamin M. Rottman^b , and Vishnu P. Murty^a

Edited by Robert Bjork, University of California Los Angeles, Los Angeles, CA; received June 30, 2023; accepted January 29, 2024

The memory benefit that arises from distributing learning over time rather than in consecutive sessions is one of the most robust effects in cognitive psychology. While prior work has mainly focused on repeated exposures to the same information, in the real world, mnemonic content is dynamic, with some pieces of information staying stable while others vary. Thus, open questions remain about the efficacy of the spacing effect in the face of variability in the mnemonic content. Here, in two experiments, we investigated the contributions of mnemonic variability and the timescale of spacing intervals, ranging from seconds to days, to long-term memory. For item memory, both mnemonic variability and spacing intervals were beneficial for memory; however, mnemonic variability was greater at shorter spacing intervals. In contrast, for associative memory, repetition rather than mnemonic variability was beneficial for memory, and spacing benefits only emerged in the absence of mnemonic variability. These results highlight a critical role for mnemonic variability and the timescale of spacing intervals in the spacing effect, bringing this classic memory paradigm into more ecologically valid contexts.

long-term memory | learning | spacing effect

While reexposing individuals to information can often enhance memory, these benefits depend on the conditions of learning. For example, repeated exposures to information that are spaced out over time benefit memory more than when repetitions are shown consecutively in a “massed” manner (1–8). While this phenomenon, known as the spacing effect, has been highly replicated (5), this work generally relies on repeated exposures to identical information. However, in the real world, we rarely encounter exactly the same information—rather our experiences contain variability, with some elements remaining stable while others change. For example, on repeat trips to your local coffee shop, your order and the coffee may be the same, but the barista serving you may vary. How does spacing affect our memories for the repeated experiences at the coffee shop? Current frameworks remain underspecified regarding how spacing influences memory in the face of mnemonic variability, limiting their ecological validity.

One theory, the “encoding variability framework,” suggests that variability is integral to the benefits associated with spacing. According to this theory, items are bound to a slowly drifting temporal context during encoding (6, 9–12). Memory traces from prior exposures are thought to be reactivated during repetitions, resulting in an integrated representation of the item with the contexts in which it was encoded (2, 6, 10, 13–17). Critically, the temporal context can then serve as a cue to facilitate recall of the associated content such that having more varied retrieval cues should enhance the probability of later memory success (6, 9–12). Since repeated items studied across longer spacing intervals (“spaced learning”) become associated with more variable temporal contexts, while consecutive repeats (“massed learning”) introduce little temporal drift and result in highly similar associated temporal contexts, memory performance should benefit more from the spaced learning conditions (1, 6, 8). However, the encoding variability hypothesis often focuses on the time-dependent drift in temporal context as the main source of variability, overlooking how variability in the content itself (i.e., different baristas) can influence memory.

Although the drift in temporal context by definition depends on the passage of time, mnemonic variability is not necessarily time dependent. For example, while the content of an experience can change as a function of time (e.g., different baristas on repeat visits), it can also remain the same (e.g., your visit always coincides with the same barista's hours). Drift in temporal context and changes in mnemonic content between repetitions thus can provide independent sources of variability, and the effects of spacing on memory may depend on the interaction of both of types of variability. When learning exposures are spaced over longer intervals, temporal and mnemonic variability may both benefit memory. However, with shorter spacing intervals, when there is little drift in temporal context,

Significance

Spacing out learning over time is one of the most robust methods to enhance later memory. Yet, much of this work is predicated on repeated exposures to identical information, which is rare in everyday life. To harness these benefits in real-world settings, we first need to characterize the efficacy of the spacing effect in the face of mnemonic variability. For isolated features of memory presented with mnemonic variability, the benefit of spacing was only observed with longer timescales of spacing intervals (hours to days) and not at short intervals (seconds to minutes). However, for more associative forms of memory, variability undermined the effects of spacing. These results challenge the existing notions of how and when spaced learning facilitates long-term memory.

Author affiliations: ^aDepartment of Psychology & Neuroscience, Temple University, Philadelphia PA 19122; and ^bDepartment of Psychology, University of Pittsburgh, Pittsburgh, PA 15260

Preprint server: <https://psyarxiv.com/xc9ut>.

Author contributions: E.T.C., Y.Z., B.M.R., and V.P.M. designed research; E.T.C. and Y.Z. performed research; E.T.C. analyzed data; and E.T.C., Y.Z., B.M.R., and V.P.M. wrote the paper.

The authors declare no competing interest.

This article is a PNAS Direct Submission.

Copyright © 2024 the Author(s). Published by PNAS. This article is distributed under [Creative Commons Attribution-NonCommercial-NoDerivatives License 4.0 \(CC BY-NC-ND\)](#).

¹To whom correspondence may be addressed. Email: emily.cowan@temple.edu.

This article contains supporting information online at <https://www.pnas.org/lookup/suppl/doi:10.1073/pnas.2311077121/-/DCSupplemental>.

Published March 12, 2024.

mnemonic variability may predominantly drive memory benefits as changes in mnemonic content may be more salient with higher fidelity in the representations of prior repetitions (18, 19). To return to our earlier example: Imagine you either spend the morning working at a coffee shop, ordering refills of your coffee, versus a coffee shop you only go to on the weekend, with relatively long intervals between visits. In the former, with shorter intervals, a change in the barista serving you may be more overt such that your memory of the experience at the coffee shop contains an integrated representation of both the stable and varying features. There is some evidence that nontemporal sources of variability benefit memory, particularly with massed learning (15, 20–22). Likewise, research on context-dependent memory has shown that variability in the environmental or incidental background context during encoding can boost memory retention (23–26). However, this work has not directly examined whether such sources of variability provide differential effects when encountered across short versus long timescales of spacing intervals, akin to how we encounter information in the real world.

In characterizing memories for more ecologically valid experiences, it is important to consider that memories can represent multiple components, including both the features of an experience as well as the associations between those features. When asked to recall their experience of a visit to a coffee shop, one may want to remember only the stable individual features (i.e., the coffee) or may also want to recall the associations between the stable and variable content (i.e., the coffee–barista combinations). However, the role of the spacing effect on item versus associative memory has rarely been considered (but see ref. 27) and has yet to be explored in the context of mnemonic variability. A secondary goal of the present studies is therefore to examine how mnemonic variability and spacing influence item versus associative memory.

To test how mnemonic variability and spacing support memory, we designed two large-scale experiments (Fig. 1*A*). To modulate mnemonic variability, participants in both experiments studied item–scene pairs where the item was either paired with four distinct scenes (variable pair) or was shown with the same scene, repeated four times (repeated pair), as shown in Fig. 1*B*.

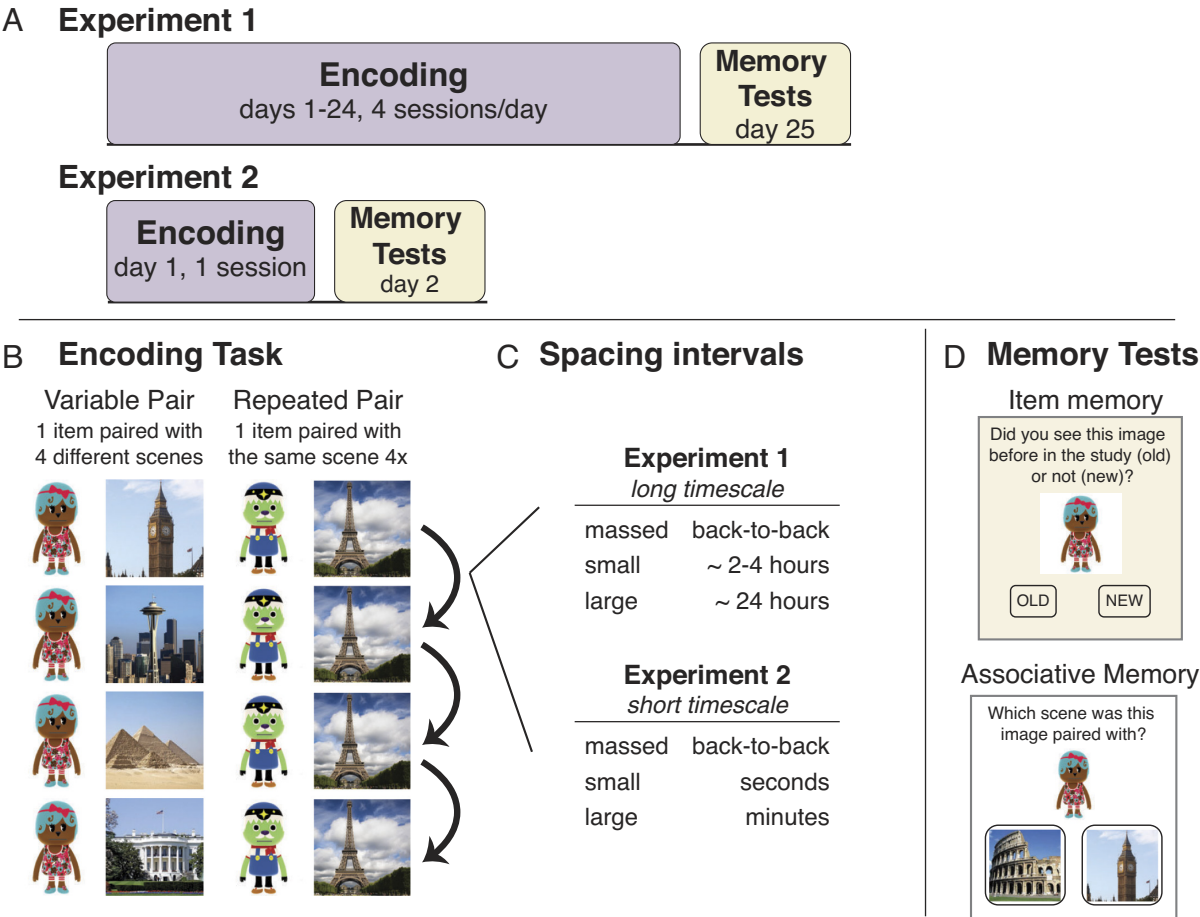


Fig. 1. Design for experiments 1 and 2. (*A*) Experiment 1 involved a 25-d protocol, with 24 d of encoding with four sessions per day, such that spacing intervals ranged from seconds to days. Item and associative memory was tested on day 25. In experiment 2, all learning was compressed into one session on 1 d, with a range of spacing intervals between seconds to minutes, and memory was tested on day 2, after a 24-h delay. (*B*) Participants were shown item–scene pairs consisting of either one item paired with four distinct scenes (variable pairs) or one item paired with the same scene that was repeated four times (repeated pairs) such that the variable pairs were associated with a source of mnemonic variability while the repeated pairs were identical across exposures. During encoding, participants were asked to imagine how the item might interact with the scene and then rate the quality (e.g., vividness) of their imagined association. (*C*) The four trials for each item–scene pairs were shown across massed, small, or large spacing intervals that ranged in relative length. In experiment 1, the spacing intervals were on a long timescale such that trials of a given pair could be viewed back to back (massed spacing intervals), shown across the four sessions within the same day (~2 to 4 h apart, “small spacing intervals”) or across four consecutive days of the experiment (24 h apart, “large spacing intervals”). In contrast, the spacing intervals in experiment 2 were compressed to a short timescale. For the four trials of a given pair, the massed spacing interval still referred to back-to-back presentation, but the small spacing intervals involved only seconds between presentations, and the large spacing intervals involved minutes between presentations. (*D*) To assess item memory, participants completed an old/new recognition task, judging whether an item was “old” and had been seen in the experiment or was new and had not been seen before. To assess associative memory for the item–scene pairs, participants were shown an item with two scenes, the paired scene and a novel foil, and asked to choose which scene the image was paired with.

In experiment 1, encoding occurred over 24 d, with 1 to 4 sessions a day, and was completed entirely on participants' mobile phones. To modulate spacing, exposures to a given pair were shown with intervening spacing intervals ranging from seconds, to hours, to days (Fig. 1C). This ecologically valid design thus mimics how repeated information can be encountered in the real world across longer timescales. To understand the role of mnemonic variability with shorter timescales of spacing intervals, in experiment 2, which was conducted on Amazon Mechanical Turk, all learning was compressed into one session on 1 d to generate shorter spacing intervals ranging only from seconds to minutes (Fig. 1C). In both experiments, item and associative memory was tested after a 24-h delay (day 25 for experiment 1 and day 2 for experiment 2; Fig. 1D). In moving away from traditional laboratory paradigms which typically test only two exposures across one spacing interval, these experiments can elucidate how mnemonic variability interacts with the benefit of spacing across multiple timescales.

Results

The structure of our experimental design allowed us to look at the influence of mnemonic variability and spacing intervals across four main questions of interest: item memory over long versus short timescales of spacing and associative memory over long versus short timescales of spacing. We first report results for item memory and then report the results from the associative memory test.

Item Memory at Long Timescales of Spacing (Experiment 1). To assess how the effects of spacing on memory are influenced by mnemonic variability with long spacing intervals, in experiment 1, we designed a 25-d study involving 24 d of encoding with four sessions per day (Fig. 1A). Participants were shown item–scene pairs consisting of either one item paired with four distinct scenes (variable pairs) or one item paired with the same scene that was repeated four times (repeated pairs, Fig. 1B). Critically, in both types of pairs, the item was always stable across the four repetitions. Thus, for the variable pairs, the association of the same item with distinct scenes introduces variability into the content itself (e.g., mnemonic variability), while the repeated pairs were always held constant across repetitions. The four trials of each pair were shown back to back within the same session (massed spacing interval), across the four sessions on the same day (small spacing interval), or across sessions on consecutive days (large spacing interval, Fig. 1C). On the last day of the study (day 25), participants completed memory tests.

Item recognition was subjected to a repeated-measures ANOVA with factors for Spacing (massed, small, and large intervals) and Pair Type (variable and repeated). We found a significant main effect of Spacing ($F(2,312) = 29.2$, $P < 0.0001$), and post hoc tests revealed that item memory was higher with longer spacing intervals, consistent with the spacing effect (massed < small: $t(313) = -5.62$, $P < 0.0001$; massed < large: $t(313) = -7.42$, $P < 0.0001$; small < large: $t(313) = -2.08$, $P = 0.039$). Turning to the effect of mnemonic variability, a significant main effect of Pair Type revealed that item memory was higher for variable pairs compared to repeated pairs ($F(1,156) = 54.23$, $P < 0.0001$).

The ANOVA also yielded a significant interaction between Spacing and Pair Type ($F(2,312) = 8.02$, $P = 0.0004$). As illustrated in Fig. 2A, while item memory was higher for the variable pairs than the repeated pairs across all spacing intervals, the difference in memory between the pair types decreased as the length of the spacing interval between exposures increased (massed: $t(156) = 6.76$, $P < 0.0001$; small: $t(156) = 2.9$, $P = 0.004$; large: $t(156) = 2.24$,

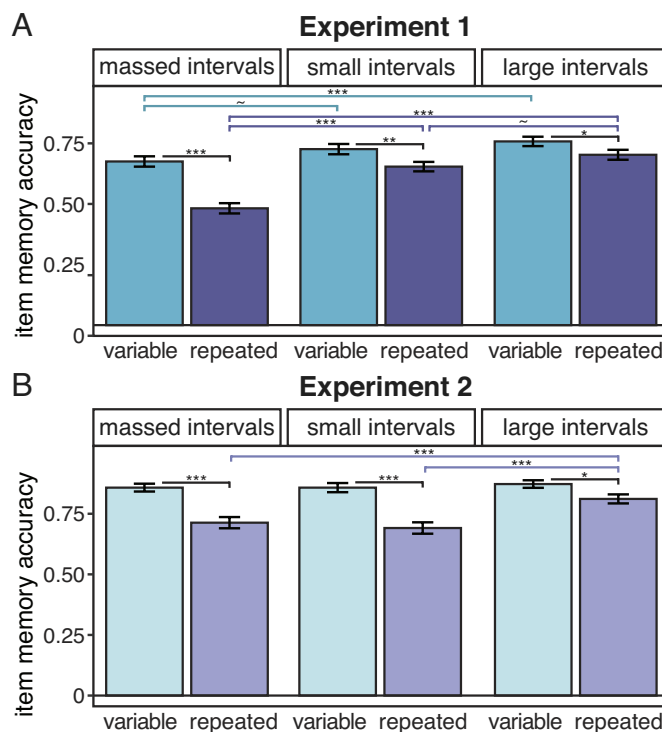


Fig. 2. Item memory in experiments 1 and 2. For item memory, significant interactions between Spacing and Pair Type indicated that mnemonic variability led to better memory than repetition when there was less spacing between exposures in both (A) experiment 1, with long spacing intervals (seconds, hours, or days), and in (B) experiment 2, with short spacing intervals (seconds or minutes). Blue and purple significance bars denote differences in item memory across spacing intervals for the variable pairs and repeated pairs, respectively. While spacing did lead to improved item memory for both the repeated and variable pairs in experiment 1, in experiment 2, only the repeated pairs showed higher item memory with longer spacing intervals. *** $P < 0.0001$; ** $P < 0.01$; * $P < 0.05$; ~ $P < 0.1$; error bars signify SEM.

$P = 0.026$). In other words, the benefit of mnemonic variability was greater when there was less spacing between exposures.

Interestingly, item memory for both the repeated and variable pairs showed benefits from spacing (Fig. 2A). For the repeated pairs, item memory was significantly higher with small and large spacing intervals compared to massed intervals (massed < small: $t(156) = -6.36$, $P < 0.00001$; massed < large: $t(156) = -7.35$, $P < 0.000001$) and marginally higher for the large compared to small spacing interval condition ($t(156) = -1.74$, $P = 0.08$). The variable pairs yielded similar results, with significantly higher item memory with large versus massed spacing intervals ($t(156) = -3.1$, $P = 0.002$) and a marginal difference between the massed versus small spacing interval conditions ($t(156) = -1.80$, $P = 0.07$), though there was not a significant difference between the small and large spacing interval conditions ($t(156) = -1.18$, $P = 0.24$). Thus, it seems that spacing exposures over intervals ranging from hours to days benefit memory for both stable and variable content.

Together, these results suggest that when repeated exposures are learned in a massed manner, items associated with mnemonic variability are better remembered than direct repetitions. However, since both variable and repeated pairs showed benefits with spacing, it seems that this pattern of results was not simply driven by a selective increase in memory for the repeated pairs at longer spacing intervals. It is possible that the temporal drift associated with longer spacing intervals may provide an independent enhancement to memory, separate from the benefits derived from changes in the content itself, such that mnemonic variability may

be less critical when repetitions are spaced over longer intervals. Importantly, these effects were not driven by differences in the delay to test across the three category blocks (*Materials and Methods* and *SI Appendix, Supplemental Information 1*).

Item Memory with Short Spacing Timescales (Experiment 2). In experiment 2, we wanted to determine whether the interaction between spacing and mnemonic variability for item memory was sensitive to the timescale of the spacing intervals. Under shorter timescales, while spacing may not provide much of a benefit, mnemonic variability could still have a beneficial effect on item memory. We modified our design by compressing the entire 24-d protocol into one 20-min session, while maintaining the relative order of trials (*Materials and Methods* and *SI Appendix, Fig. S1*). Thus, in experiment 2, the relative massed, small, and large spacing intervals only ranged from seconds to minutes, with pairs shown back to back or with intervening trials between exposures (Fig. 1C).

As above, item recognition was analyzed using a repeated-measures ANOVA with factors for Spacing (massed, small, and large intervals) and Pair Type (variable and repeated). There was again a significant main effect of Spacing ($F(2,270) = 7.70$, $P = 0.0006$). Item memory for pairs shown across large spacing intervals (e.g., with intervening items) was significantly greater than for pairs shown with massed ($t(271) = -3.08$, $P = 0.0023$) or small spacing intervals ($t(271) = -3.69$, $P = 0.0003$). However, in experiment 2, there was no significant difference between item memory for the massed and small spacing interval conditions ($t(271) = 0.58$, $P = 0.56$), perhaps due to the fact that in the compressed experiment 2, pairs in both conditions were shown essentially back to back, whereas in experiment 1, these trials were separated by approximately 2 h (*Materials and Methods*).

We also found a significant main effect of Pair Type such that variable pairs were remembered better than repeated pairs ($F(1,135) = 73.01$, $P < 0.0001$), as well as a significant Spacing $\times$ Pair Type interaction effect ($F(2,270) = 4.58$, $P = 0.011$). Consistent with the results from experiment 1, the benefit of the variable over repeated pairs decreased as the spacing interval increased (massed: $t(135) = 5.67$, $P < 0.0001$; small: $t(135) = 6.13$, $P < 0.0001$; large: $t(135) = 2.51$, $P = 0.013$, Fig. 2B). These results provide further evidence suggesting that under massed conditions, mnemonic variability benefits item memory more than direct repetitions.

Interestingly, in experiment 2, only the repeated pairs showed a benefit from spacing (Fig. 2B). Item memory was significantly higher for the large spacing interval conditions compared to both the massed and small intervals (massed $<$ large: $t(135) = -3.37$, $P = 0.001$; small $<$ large: $t(135) = -4.26$, $P < 0.0001$), although memory in the massed and small spacing interval conditions did not significantly differ ($t(135) = 0.77$, $P = 0.44$). In contrast, for the variable pairs, there were no significant differences across any of the spacing interval conditions (massed vs. small: $t(135) < 0.01$, $P = 1$; massed vs. large: $t(135) = -0.68$, $P = 0.50$; small vs. large: $t(135) = -0.65$, $P = 0.51$). Thus, unlike in experiment 1, it seems that when all spacing is compressed across shorter timescales, and thus generally involves more massed learning, spacing only provided additional benefits for the repeated pairs and did not modulate item memory for the variable pairs.

Associative Memory across Long Timescales of Spacing (Experiment 1). Next, we examined associative memory to understand how mnemonic variability and spacing influence memory for the links between item and scenes that either remained

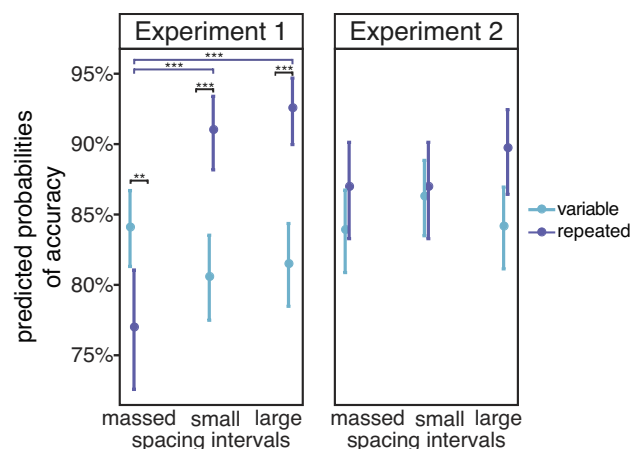


Fig. 3. Associative memory in experiments 1 and 2. In experiment 1 (*Left*), repeated pairs were remembered better than the variable pairs with small or large spacing intervals, but mnemonic variability led to enhanced memory compared to variable pairs when presented in a massed manner. In experiment 2 (*Right*), with shorter timescales of spacing intervals, the repeated pairs were generally better remembered than the variable pairs, but spacing did not benefit memory for either pair type. The plot represents the predicted probabilities from the model fit; error bars indicate 95% CI. *** $P < 0.0001$; ** $P < 0.01$.

stable or changed over repetitions. We first examined associative memory in experiment 1, when spacing intervals ranged from seconds to days, using a logistic mixed-effects model with fixed effects for the interaction between Spacing (massed, small, and large intervals) and Pair Type (variable and repeated) and a random effect of subject. This analysis yielded a significant main effect of Spacing ($X^2 = 7.09$, $P = 0.029$); in line with the spacing effect, overall associative memory was higher for pairs learned with spacing between trial presentations compared to back to back, within the same session (massed $<$ small: $\beta = -0.43$, $P = 0.0003$; massed $<$ large: $\beta = -0.56$, $P < 0.0001$), although there was not a significant difference between the small and large spacing interval conditions ($\beta = -0.13$, $P = 0.55$). This analysis also yielded a significant main effect of Pair Type ($X^2 = 11.60$, $P = 0.0007$), with higher associative memory for the repeated pairs compared to the variable pairs.

We also found a significant interaction between Spacing and Pair Type ($X^2 = 56.06$, $P < 0.0001$). While the main effect of Pair Type revealed higher associative memory for the repeated pairs compared to the variable pairs overall, post hoc tests suggest this pattern differed across spacing conditions (Fig. 3). For the massed spacing intervals, associative memory for the variable pairs was greater than for repeated pairs ($\beta = 0.45$, $P = 0.001$), but the opposite pattern was observed for both the small and large interval conditions such that repeated pairs were significantly better remembered than the variable pairs (small: $\beta = -0.89$, $P < 0.0001$; large: $\beta = -1.04$, $P < 0.0001$).

Further, only the repeated pairs seemed to show a significant increase in associative memory with spacing. For the repeated pairs, associative memory was higher with longer spacing intervals between exposures (massed $<$ small: $\beta = -1.10$, $P < 0.0001$; massed $<$ large: $\beta = -1.31$, $P < 0.0001$), although there was no difference between the small and large spacing interval conditions ($\beta = -0.21$, $P = 0.63$). In contrast, for the variable pairs, there were no significant differences in associative memory across any of the three spacing intervals (massed vs. small: $\beta = 0.24$, $P = 0.11$; massed vs. large: $\beta = 0.18$, $P = 0.29$; small vs. large: $\beta = -0.06$, $P = 0.87$).

These results demonstrate that associative memory for the repeated pairs benefits more from longer intervals between

learning exposures than the variable pairs, but when learning is more condensed, mnemonic variability may enhance memory over repeated pairs. As above, these effects were not driven by differences in the delay to test for the three category blocks (*SI Appendix, Supplemental Information 1*).

Associative Memory with Short Timescales of Spacing (Experiment 2). We next examined associative memory for experiment 2, in which learning was completed during a single session, to determine 1) how time contributes to the observed spacing benefit for repeated pairs and 2) how mnemonic variability influences associative memory under more generally compressed learning conditions.

A logistic mixed effects model with fixed effects for the interaction between interaction between Spacing (massed, small, and large) and Pair Type (variable and repeated) and a random effect of subject yielded a significant main effect of Pair Type ($X^2 = 6.94$, $P = 0.008$); associative memory was higher for the repeated pairs compared to the variable pairs (Fig. 3). Unlike in experiment 1, there was not a significant main effect of Spacing ($X^2 = 1.36$, $P = 0.51$) or a significant interaction between Pair Type and Spacing ($X^2 = 3.16$, $P = 0.21$).

These results suggest that repetitions of stable information may be generally more beneficial to associative memory than variability. However, while in experiment 1, associative memory for the repeated pairs showed a benefit from spacing over longer time scales, there was no such effect in experiment 2. In addition, for the variable pairs, the mechanisms of the spacing effect seem to fail at both short and long timescales of spacing intervals, perhaps owing to an inability in integrating across the pairs or interference across the scene associates in the pairs (*SI Appendix, Supplemental information 2*).

Discussion

In the current set of experiments, we examined how mnemonic variability interacts with the spacing effect over distinct timescales to influence item and associative memory. With regard to item memory, in both experiments, we found that mnemonic variability led to greater benefits than direct repetitions, particularly when exposures were shown in a massed fashion. For items paired with distinct scenes on each exposure (i.e., variable pairs), the memory benefit from spacing out repeated learning was only observed with longer timescales of spacing intervals (hours to days); we did not find a benefit of spacing with short timescales (seconds to minutes). In contrast, items paired with the same scene (i.e., repeated pairs) showed a consistent spacing benefit across both short and long spacing interval timescales. For associative memory, a different pattern emerged: Memory was higher for the repeated pairs overall, and only those pairs showed a spacing benefit with long timescales, while associative memory for the variable pairs was not modulated by the spacing interval in either experiment. Together, these results highlight the importance of considering the mnemonic variability present in the learned content for understanding how learning conditions affect memory. These results suggest that the benefits of spacing are not absolute but instead seem to be modulated by key factors including the variability present in the mnemonic content across repetitions, the timescale of the spacing interval between exposures, and the type of memory measure.

Considerable prior work has demonstrated that spacing repeated learning opportunities benefits memory more than massed learning conditions (1–7), an effect that has been attributed to the infusion of variability into the encoding experience through the time-dependent drift in temporal context (6, 9, 10, 12).

However, this research has often overlooked the way we encounter repeated information in the real world, in which content may be only partially overlapping, with some elements remaining stable and some varying over experiences, in a manner that is not necessarily time dependent. The present results provide evidence that the benefits of spacing may depend on both sources of variability—from changes in temporal drift and in the content itself—depending on the timescale of the spacing interval.

Mnemonic variability benefited item memory more than direct repetitions, but this advantage decreased with longer spacing intervals between exposures. As such, it seems that with shorter spacing intervals or massed learning, changes in the content of the memoranda over repetitions may be more important for later item memory. These results are generally consistent with the limited prior work examining the relationship between nontemporal forms of variability and the spacing effect, which found that under massed conditions, mnemonic variability resulted in better memory than identical repetitions and that mnemonic variability resulted in a reduced spacing effect compared to direct repetitions (15, 20–22). When learned in closer succession, memory for the prior exposures may be more precise, engendering a greater novelty signal in response to the variation in the content (18, 19), which in turn can strengthen memory (28–30). This may also explain why spacing did not modulate item memory for the variable pairs in experiment 2; with more compressed learning conditions, the benefits of mnemonic variability may provide a sufficient boost to item memory, overshadowing the benefits from small changes in temporal drift available over such short intervals. At the same time, for the repeated pairs, these limited changes in temporal drift could still have facilitated the observed spacing benefit in the absence of a source of mnemonic variability. In contrast, with the longer timescales in experiment 1, both the repeated and variable pairs showed greater memory with increasing spacing between learning exposures. In this case, temporal drift could either outweigh mnemonic variability or there could be an additive benefit from both sources of variability on item memory.

For associative memory, overall, the repeated pairs were better remembered than the variable pairs in both experiments. Further, associative memory for the variable pairs was not modulated by the spacing intervals at either timescale. It therefore seems that associative memory does not benefit from mnemonic variability in the same way as item memory. It is possible that for item memory, mnemonic variability may have served to strengthen the item's representation in memory, generating multiple retrieval cues centered on the item itself, due to the pairing with four distinct scenes. On the other hand, for associative memory, mnemonic variability may have resulted in four distinct item–scene memory traces, which might be weaker compared to an integrated trace. Prior work has shown that the neural representations for memories with overlapping content can become differentiated with learning (31–35). Such “repulsion” processes, which can help reduce interference across distinct memories (33, 34, 36), could actually inhibit the benefits of spacing. To produce a spacing effect, content must be recognized as a repeat of previously studied information (2, 6). If the variable pairs undergo such a repulsion process and come to be represented by independent, differentiated associative memories, it may reduce the ability to reactivate and integrate across learning exposures. Thus, associative memory for the repeated pairs showed a spacing benefit over long timescales, but the variable pairs did not. Considering that traditionally spacing effects are based on identical repetitions, here we show evidence that for associative memory, the benefits of spacing do not generalize to partially overlapping content.

While this pattern suggests differences in the effects of mnemonic variability on item and associative memory, there is a possibility that responses to the associative memory test could have been based on scene recognition, as the trials involved an item cue with one old scene and one novel foil. However, in this case, we would have expected to find a similar pattern of results for both memory tests, at least for the repeated pairs, where the scenes and items are seen with equal frequency. Notably, we used this design to reduce further learning episodes during testing, given our focus on equally manipulating repetitions across conditions, but future work using a test comparing intact versus rearranged pairings could confirm this dissociation between types of memory measures. Further, one puzzling finding from the associative memory test is that while there was a similar benefit of mnemonic variability over repetition when pairs were shown in a consecutive, massed fashion in experiment 1, this result was not apparent in experiment 2, with the shorter timescale of spacing intervals. This discrepancy raises an interesting question about how the relative volatility of the learning conditions impacts the way associative information is processed. In the compressed timescale of experiment 2, when there is generally more potential for interference, the stability of massed repetition of the same content may facilitate a boost in memory, while in experiment 1, presenting only the massed repetitions in a given encoding session could lead to poor memory. Future studies could more directly examine how the volatility of the general learning context and cognitive load impact the way massed information is remembered.

Interestingly, for associative memory, the repeated pairs only showed a spacing benefit when the intervals between exposures lasted hours or days, and not when the exposures were repeated over a period of seconds or minutes, suggesting that such effects may depend on the passage of time itself. This may be in line with an alternative account for the spacing effect, which proposes that spacing benefits memory by providing an opportunity for processes of offline memory consolidation to stabilize new memories (27, 37). Since the hippocampus is critical in both consolidation-related processes (38–41) and associative memory (42–44), such an account could fit with the distinct role of mnemonic variability in item versus associative memory. Alternatively, these results may also fit with the retrieval practice account of the spacing effect, which posits that spacing benefits memory by reducing the accessibility of learned content, resulting in more difficult retrieval processes when reexposed to repetitions (45). As applied to the current data, compared to the short timescale of spacing intervals in experiment 2, the longer spacing intervals in experiment 1 could lead to more difficult retrieval processes for the repeated pairs. However, this account does not seem able to explain the dissociation in item and associative memory for the repeated pairs in experiment 2; if retrieval practice is not sufficiently engaged to benefit associative memory, it is unclear why item memory for the repeated pairs would still show a spacing benefit. Additionally, this account might suggest that the variable pairs do not trigger retrieval practice mechanisms and thus fail to show benefits of spacing for associative memory, but again, it is difficult to reconcile this explanation with the observed significant spacing effect for item memory in experiment 1. Future work could directly examine the role of retrieval practice in item and associative memory by encouraging participants to actively retrieve stimuli across encoding sessions.

In addition to the timescale of spacing intervals in experiments 1 and 2, it is also important to consider that other factors may differ between the two experiments. While experiment 1 was conducted on participants' cell phones over the course of 25 d, experiment 2 was conducted on Amazon Mechanical Turk in one

sitting. As a result, the potential for distraction during the task could be greater in experiment 1, while the cognitive load and potential for fatigue may be greater in experiment 2. However, it is important to consider that factors like increased cognitive load and fatigue are relatively inherent with shorter timescales of learning. In addition, there were some differences in the demographics of the two samples (mean age: experiment 1 = 21; experiment 2 = 44), which could have contributed to the different effects on memory performance across the two studies. While the main analyses were within-subjects comparisons of the variable versus the repeated pairs, future studies could reduce such sample-based differences to draw more direct cross-timescale dissociations.

Together, the present experiments represent large sample investigations into the effects of an ecologically valid source of mnemonic variability on the spacing effect. Our work differs from standard laboratory paradigms testing the spacing effect, which often involve only one repetition of the same exact stimuli. Indeed, conducting experiment 1 entirely on participants' cell phones, with single trial exposures to repeated or variable stimuli over spacing intervals lasting seconds, hours, or days, represents a more naturalistic design akin to how we actually encounter overlapping or repeated information in the world. This work therefore extends current frameworks of the spacing effect, demonstrating interactions between the source of variability and the timescale of spacing intervals that influence item and associative memory.

Materials and Methods

Both experiments 1 and 2 were preregistered. The preregistration for experiment 1 can be found at <https://osf.io/jk7qr>. The preregistration for experiment 2 can be found at <https://osf.io/wem4v>.

Participants.

Experiment 1. A total number of 311 participants were recruited for experiment 1. Participants who missed more than 8 encoding sessions were not allowed to complete the experiment. Participants were excluded from analysis if they missed more than 1 repetition and 1 variable trial and based on memory performance if 1) overall item memory is below chance and 2) overall associative memory is below chance for all variable pairs and all repetition pairs in each category. A total of 102 participants were excluded for not completing a sufficient number of encoding trials, and 47 participants were excluded based on memory performance. A further five participants were excluded because of technical issues with data collection, leaving a total of 157 participants (mean age = 21; 45 male, 111 female, and 1 nonbinary/other) included in all experiment 1 analyses. This protocol was reviewed and approved by the IRB at the University of Pittsburgh under the Benign Behavioral intervention (low risk) exempt category. To ensure that participants made an informed decision to participate, they were guided through an introductory script that described the research study and explained that it is voluntary.

Experiment 2. Participants completed experiment 2 via Amazon Mechanical Turk and were provided a monetary bonus for completing both sessions. Similar exclusion criteria were applied in experiment 2, adapted to the 2-d vs. 25-d protocol. Participants were excluded from analyses if they did not complete both days of the experiment or showed poor memory performance as detailed above. One hundred thirty-nine participants out of 171 total recruited completed the full experiment (e.g., both days). Three participants were excluded based on memory performance, for a total of 136 participants (mean age = 44; 69 male, 66 female, and 1 prefer not to disclose) included in all experiment 2 analyses. Notably, the sample size for experiment 2 was based on a power analysis conducted on experiment 1 (see preregistration). This protocol was reviewed and approved by the IRB at Temple University, and informed consent was obtained from all participants at the start of the first experimental session.

Procedure.

Experiment 1. This study involved a 25-d protocol, consisting of a 24-d learning phase with four sessions per day and a test phase on day 25 (Fig. 1A). The entire experiment was completed on participants' mobile phones. This experiment was

part of a larger study, with other tasks, including a causal learning study and reinforcement learning study, occurring during the 25-d protocol. Stimulus content across tasks was controlled to avoid interference. During a Zoom meeting on the first day of the study (day 0), participants were provided instructions about the study and complete practice sessions to increase familiarity with the stimuli and procedure of the learning phase. On days 1 to 24 of the experiment, the four sessions per day occurred at 9 am, 11 am, 1 pm, and between 3 and 9 pm, accommodating the requirements for the designs of the other tasks (*SI Appendix, Fig. S1A*). For each session, participants had a 2-h time window to complete the task. Participants received text message reminders about their participation every 30 min within this window until they completed the session. On day 25, participants completed two memory tests via their cell phone web browser. At the time of each session, participants received text messages providing a link to follow to complete that session's tasks, followed by reminders every 30 min until participants completed the task. If the task was not completed within the 2-h window, the session was counted as a missed session; if a participant missed more than eight sessions, they were not allowed to continue with the study. Participants were provided a monetary bonus for completing the entire 25-d protocol.

Experiment 2. This study involved a 2-d protocol completed via Amazon Mechanical Turk (Fig. 1A). The session on day 1 consisted of all the encoding trials, while participants completed the memory tests on day 2. Participants were provided a monetary bonus for completion of both sessions.

Stimuli. The same stimuli pool was utilized in both experiments 1 and 2. Item-scene pairs were generated from three possible categories consisting of cartoon characters paired with famous scenes, tools paired with indoor scenes, or animals and cartoon scenes. We chose to utilize three different categories of stimuli in a blocked fashion to maximize the number of trials in each condition of interest while minimizing interference across the long timescale of the experiment. Cartoon scenes were generated from the Adobe Stock Free collection but customized to remove any additional figures or features. All other stimuli were from in-lab databases. Across the encoding and memory test sessions, a total of 33 item stimuli and 36 scene stimuli were included from each category. Two lists of item-scene pairs were randomly generated and counterbalanced across participants.

Encoding Task.

Experiment 1. In the encoding task, participants saw item-scene pairs and were asked to "imagine how the image would interact with the scene." Participants then rated the "vividness of your imagined association" on a scale of 1 to 4, where 1 meant "poor and difficult to imagine," and 4 meant "a vivid imagined association," by clicking the number on their phone's screen. Pairs were shown for 5 s, and the vividness judgments were self-paced. There were two types of pairs: variable or repetition. Both pair types consist of 4 item-scene trials; however, for the variable pairs, one item was paired with four distinct scenes, each shown once, while the repetition pairs consist of one item and one scene repeated four times (Fig. 1B).

To modulate spacing, the four variable/repetition trials could be encountered back to back in the same session ("massed spacing intervals"), shown across the four sessions within the same day (one trial per session, "small spacing intervals"), or across four consecutive days of the experiment (one trial per session, shown in the same session block on each day to match time of day, "large spacing intervals"). Therefore, the massed trials were presented with only seconds between, while the trials in the small spacing intervals condition were shown with an intervening ~2-h delay between exposures, and the trials with large spacing intervals had approximately 24 h between exposures (Fig. 1C). Additional filler pairs consisting of a unique item and scene were shown only once (i.e., neither the item nor scene is repeated). These filler pairs were presented either within the same session (four pairs in a row) or spread across days and sessions (one pair per session), mimicking the psychological experience of the variable and repetition trials of interest. *SI Appendix, Fig. S1A* provides an example illustration of the distribution of trials in each spacing condition. Three categories of pairs were shown across the 24-d protocol in a blocked manner, with eight consecutive days of trials shown in the same category. The order in which categories were shown was randomly assigned across participants.

In total, the encoding task involved 96 encoding trials, with three variable and three repetition pairs in each stimulus category (e.g., 1 pair in each spacing and stimulus condition in each category, for a total of nine pairs each) and 24 total filler pairs across the three categories. The assignment of pair and spacing conditions to the encoding sessions (e.g., session/day) was pseudorandomized

to optimally differentiate the order of the pair conditions and spacing. The same order was repeated for each of the three category blocks. Three possible orders were generated, and participant assignment was counterbalanced.

Experiment 2. Experiment 2 used the same protocol as experiment 1 except for the following changes. First, all encoding was compressed into one session on 1 d. The same absolute order of trials was retained from experiment 1, thus eliminating the time delays between sessions. Therefore, for the massed spacing intervals, the four trials for each variable/repetition pair were still presented back to back, while the pairs in the small and large spacing interval conditions were shown with varying numbers of intervening trials from other pairs (Fig. 1C). An example of the trial distribution in experiment 2 is provided in *SI Appendix, Fig. S1B*. Second, in experiment 1, participants were required to click a link to begin each encoding session. To maintain this source of agency in experiment 2, following the vividness judgment, participants were asked to click a button to begin the next trial. Critically, this button only appeared for trials in the small or large interval conditions and did not interrupt the consecutive nature of the massed presentation. Participants had 5 s to make their vividness judgments; once participants made a response, the task moved on.

Memory Tests. On day 25 (experiment 1) or day 2 (experiment 2), participants completed two memory tests (Fig. 1D). In the item recognition task, participants were instructed to judge whether an item was old and had been seen in the experiment or was new and had not been seen before in the experiment. All 42 items from encoding were included (14 per category), intermixed with 21 novel foils (7 from each category). Following the item recognition task, participants completed an associative memory test on the item-scene pairings. An item from encoding was shown above two scene images: One scene was paired with the item during the encoding task, and the other scene was a novel foil from the same category. We chose to use a novel foil in this task, rather than showing two old scenes as options to reduce interference across the small number of trials and dissuade retrieval strategies like recall to reject. Participants had to select which scene had been shown together with the item cue during encoding. We tested memory on all of the repetition pairs and both the first and last paired scenes for the variable pairs, as well as half of the filler pairs. We chose not to test every associate in order to 1) limit the number of test trials in this complex design, particularly for experiment 1, and 2) due to concerns that repeated testing could inflate memory, potentially facilitating the reactivation of other scenes associated with the same cue. Including the first and last scene associates thus allowed us to test for interference effects across variable pair associates while balancing these concerns. In experiment 1, participants' responses for both memory tests were made by clicking directly on the buttons on their touchscreen and were self-paced; in experiment 2, responses were made using their computer's keyboard and had an upper limit of 10 s before the trial advanced. For each memory test, the order of trials was counterbalanced from two generated random orders corresponding to the item-stimulus pair list assigned during encoding.

Analysis. All analyses were performed in MATLAB and RStudio. We used R packages including lme4, eemmeans, rstatix, sjPlot, and ggplot2. Accuracy on the item memory test was calculated as the proportion of items correctly identified as old out of the total number of trials. Results within each experiment were analyzed using repeated-measures ANOVAs, and follow-up repeated measures *t* tests were used when appropriate. For the associative memory task, overall accuracy for the exclusion criterion was measured as the proportion of trials participants correctly selected the scene that had been paired with the cue item out of the total number of trials. For analyses of associative memory, we used logistic mixed-effects models (with glmer) with a binary outcome variable indicating correct versus incorrect accuracy on each trial. Using logistic mixed effects models better maximized data inclusion considering our exclusion criterion that allowed participants to miss 1 variable and repeat trial such that they may not have seen the given item-scene association that was tested. We first tested a model comparison of a null model including the main effects for fixed effects of Pair Type (variable and repeated) and Spacing (massed, small, and large intervals) and a random effect of subject, versus a model with the interaction of these fixed effects. Inclusion of additional random slopes terms led these models to become singular, indicating a lack of sufficient variance to warrant inclusion of the random slopes. We tested for significance using type

II Wald Chi Square tests within the Anova() command. Significant effects were examined with post hoc tests using the eemmeans package.

Data, Materials, and Software Availability. Anonymized behavioral data have been deposited in OSF (<https://osf.io/k2yg9/>) (46).

1. A. W. Melton, The situation with respect to the spacing of repetitions and memory. *J. Verbal Learn. Verbal Behav.* **9**, 596–606 (1970).
2. D. L. Hintzman, J. J. Summers, R. A. Block, What causes the spacing effect? Some effects of repetition, duration, and spacing on memory for pictures. *Memory Cogn.* **3**, 287–294 (1975).
3. F. N. Dempster, Spacing effects and their implications for theory and practice. *Educ. Psychol. Rev.* **1**, 309–330 (1989).
4. F. N. Dempster, "Chapter 9-Distributing and managing the conditions of encoding and practice" in *Memory*, E. L. Bjork, R. A. Bjork, Eds. (Academic Press, 1996), pp. 317–344.
5. N. J. Cepeda, H. Pashler, E. Vul, J. T. Wixted, D. Rohrer, Distributed practice in verbal recall tasks: A review and quantitative synthesis. *Psychol. Bull.* **132**, 354–380 (2006).
6. G. B. Maddox, Understanding the underlying mechanism of the spacing effect in verbal learning: A case for encoding variability and study-phase retrieval. *J. Cogn. Psychol.* **28**, 684–706 (2016).
7. A. M. Glenberg, Monotonic and nonmonotonic lag effects in paired-associate and recognition memory paradigms. *J. Verbal Learn. Verbal Behav.* **15**, 1–16 (1976).
8. D. Hintzman, "Theoretical implications of the spacing effect" in *Theories in Cognitive Psychology: The Loyola Symposium*, R. L. Solso, Ed. (Lawrence Erlbaum Associates, Potomac, MD, 1974), pp. 77–99.
9. W. K. Estes, Statistical theory of distributional phenomena in learning. *Psychol. Rev.* **62**, 369–377 (1955).
10. A. M. Glenberg, Component-levels theory of the effects of spacing of repetitions on recall and recognition. *Memory Cogn.* **7**, 95–112 (1979).
11. M. W. Howard, M. J. Kahana, A distributed representation of temporal context. *J. Math. Psychol.* **46**, 269–299 (2002).
12. A. S. Benjamin, J. Tullis, What makes distributed practice effective? *Cogn. Psychol.* **61**, 228–247 (2010).
13. L. L. Siegel, M. J. Kahana, A retrieved context account of spacing and repetition effects in free recall. *J. Exp. Psychol. Learn. Mem. Cogn.* **40**, 755–764 (2014).
14. R. L. Greene, Spacing effects in memory: Evidence for a two-process account. *J. Exp. Psychol. Learn. Memory Cogn.* **15**, 371–377 (1989).
15. P. P. J. L. Verhoeven, R. M. J. P. Rikers, H. G. Schmidt, Detrimental influence of contextual change on spacing effects in free recall. *J. Exp. Psychol. Learn. Memory Cogn.* **30**, 796–800 (2004).
16. S. J. Thios, P. R. D'Agostino, Effects of repetition as a function of study-phase retrieval. *J. Verbal Learn. Verbal Behav.* **15**, 529–536 (1976).
17. N. Kornell, R. A. Bjork, M. A. Garcia, Why tests appear to prevent forgetting: A distribution-based bifurcation model. *J. Memory Lang.* **65**, 85–97 (2011).
18. D. Kumaran, E. A. Maguire, An unexpected sequence of events: Mismatch detection in the human hippocampus. *PLoS Biol.* **4**, e424 (2006).
19. D. Kumaran, E. A. Maguire, Match-mismatch processes underlie human hippocampal responses to associative novelty. *J. Neurosci.* **27**, 8517–8524 (2007).
20. D. Dellarosa, L. E. Bourne, Surface form and the spacing effect. *Memory Cogn.* **13**, 529–537 (1985).
21. J. A. Glover, A. J. Corkill, Influence of paraphrased repetitions on the spacing effect. *J. Educat. Psychol.* **79**, 198–199 (1987).
22. S. L. Appleton-Knapp, R. A. Bjork, T. D. Wickens, Examining the spacing effect in advertising: Encoding variability, retrieval processes, and their interaction. *J. Consumer Res.* **32**, 266–276 (2005).
23. S. M. Smith, J. D. Handy, Effects of varied and constant environmental contexts on acquisition and retention. *J. Exp. Psychol. Learn. Memory Cogn.* **40**, 1582–1593 (2014).
24. S. M. Smith, J. D. Handy, The crutch of context-dependency: Effects of contextual support and constancy on acquisition and retention. *Memory* **24**, 1134–1141 (2016).
25. S. M. Smith, E. Z. Rothkopf, Contextual enrichment and distribution of practice in the classroom. *Cogn. Instruct.* **1**, 341–358 (1984).
26. S. M. Smith, A. Glenberg, R. A. Bjork, Environmental context and human memory. *Memory Cogn.* **6**, 342–353 (1978).
27. L. Litman, L. Davachi, Distributed learning enhances relational memory consolidation. *Learn. Memory* **15**, 711–716 (2008).
28. J. E. Lisman, A. A. Grace, The hippocampal-VTA loop: Controlling the entry of information into long-term memory. *Neuron* **46**, 703–713 (2005).
29. C. Ranganath, G. Rainer, Cognitive neuroscience: Neural mechanisms for detecting and remembering novel events. *Nat. Rev. Neurosci.* **4**, 193–202 (2003).
30. D. Shohamy, R. A. Adcock, Dopamine and adaptive memory. *Trends Cogn. Sci.* **14**, 464–472 (2010).
31. G. Wanjia, S. E. Favila, G. Kim, R. J. Molitor, B. A. Kuhl, Abrupt hippocampal remapping signals prevents interference from memory. *Nat. Commun.* **12**, 4816 (2021).
32. A. J. H. Chanale, A. Oza, S. E. Favila, B. A. Kuhl, Overlap among spatial memories triggers repulsion of hippocampal representations. *Curr. Biol.* **27**, 2307–2317.e5 (2017).
33. S. E. Favila, A. J. H. Chanale, B. A. Kuhl, Experience-dependent hippocampal pattern differentiation prevents interference during subsequent learning. *Nat. Commun.* **7**, 11066 (2016).
34. J. C. Hulbert, K. A. Norman, Neural differentiation tracks improved recall of competing memories following interleaved study and retrieval practice. *Cereb. Cortex* **25**, 3994–4008 (2015).
35. H. R. Dimsdale-Zucker, M. Ritchey, A. D. Ekstrom, A. P. Yonelinas, C. Ranganath, CA1 and CA3 differentially support spontaneous retrieval of episodic contexts within human hippocampal subfields. *Nat. Commun.* **9**, 294 (2018).
36. A. J. H. Chanale, A. G. Tremblay-McGaw, M. L. Drascher, B. A. Kuhl, Adaptive repulsion of long-term memory representations is triggered by event similarity. *Psychol. Sci.* **32**, 705–720 (2021).
37. K. L. Vilberg, L. Davachi, Perirhinal-hippocampal connectivity during reactivation is a marker for object-based memory consolidation. *Neuron* **79**, 1232–1242 (2013).
38. J. L. McClelland, B. L. McNaughton, R. C. O'Reilly, Why there are complementary learning systems in the hippocampus and neo-cortex: Insights from the successes and failures of connectionists models of learning and memory. *Psychol. Rev.* **102**, 419–457 (1995).
39. E. T. Cowan, A. C. Schapiro, J. E. Dunsmoor, V. P. Murty, Memory consolidation as an adaptive process. *Psychonomic Bull. Rev.* **28**, 1796–1810 (2021).
40. M. Moscovitch, R. Cabeza, G. Winocur, L. Nadel, Episodic memory and beyond: The Hippocampus and Neocortex in transformation. *Ann. Rev. Psychol.* **67**, 105–134 (2016).
41. J. Robin, M. Moscovitch, Details, gist and schema: Hippocampal-neocortical interactions underlying recent and remote episodic and spatial memory. *Curr. Opin. Behav. Sci.* **17**, 114–123 (2017).
42. L. Davachi, Item, context and relational episodic encoding in humans. *Curr. Opin. Neurobiol.* **16**, 693–700 (2006).
43. C. Ranganath, A unified framework for the functional organization of the medial temporal lobes and the phenomenology of episodic memory. *Hippocampus* **20**, 1263–1290 (2010).
44. H. Eichenbaum, A. P. Yonelinas, C. Ranganath, The medial temporal lobe and recognition memory. *Ann. Rev. Neurosci.* **30**, 123–152 (2007).
45. R. A. Bjork, E. L. Bjork, "A new theory of disuse and an old theory of stimulus fluctuation" in *From Learning Processes to Cognitive Processes: Essays in Honor of William K Estes*, A. Healy, S. Kosslyn, R. Shiffrin, Eds. (Erlbaum, Hillsdale, NJ, 1992), vol. 2, pp. 35–67.
46. E. T. Cowan, Y. Zhang, B. M. Rottman, V. Murty, The effects of mnemonic variability and spacing on memory over multiple timescales. OSF. <https://osf.io/k2yg9/>. Deposited 20 February 2024.

ACKNOWLEDGMENTS. We are grateful to Sarah DuBrow and Benjamin Hutchinson for helpful discussions on the design of this study. This work was supported by NSF grant 1651330 to B.M.R. and NIH R21 DA043568, K01 MH111991, and R01 DA055259 to V.P.M.