Optimizing Learning Efficiency: Balancing Spacing and Repetition Under Time Constraints

Meng Cao (mengc2@andrew.cmu.edu)
Human-Computer Interaction Institute, Carnegie Mellon University, 5000 Forbes Ave
Pittsburgh, PA 15213 USA

Veronica X. Yan (veronicayan@austin.utexas.edu)
Department of Educational Psychology, The University of Texas at Austin, 1912 Speedway M/S D5800
Austin, TX 78712, USA

Faria Sana (fsana@athabascau.ca)
Center for Social Sciences, Athabasca University, 10011 109 Street Northwest, #1200
Edmonton, Alberta T5J 3S8, Canada

Paulo F. Carvalho (pcarvalh@andrew.cmu.edu)
Human-Computer Interaction Institute, Carnegie Mellon University, 5000 Forbes Ave
Pittsburgh, PA 15213 USA

Abstract

Spaced retrieval practice has been repeatedly demonstrated to improve learning, but its implementation is often constrained by real-world time limitations. This study investigated whether, under fixed study durations, learners should prioritize spacing or repetition. Across two experiments (total N = 1589), participants practiced Indonesian vocabulary under four conditions that varied in spacing and repetition. Item difficulty was also manipulated. Results showed that increasing repetitions at the cost of spacing enhanced immediate test performance, particularly for harder items. These findings suggest that spaced retrieval practice is effective only when learners have sufficient prior repetitions to retrieve information successfully. This study highlights the trade-offs between spacing and repetition under time constraints and offers practice insights for optimizing learning strategies.

Keywords: spacing, retrieval practice, optimization, repeated practice, vocabulary learning

Introduction

In cognitive science, spacing (or distributed practice) and retrieval practice (or testing) are two of the most robust and widely studied cognitive principles for enhancing learning (e.g., Carpenter et al., 2022; Dunlosky et al., 2013; Pashler et al., 2007; Yan et al., 2024). Combining the two principles, spaced retrieval practice has emerged as a particularly effective strategy for promoting durable learning (for a review, see Latimier et al., 2021). This approach spaces retrieval practice sessions over time, optimizing both memory strength and retention.

However, implementing spaced retrieval practice in real-world educational settings poses some challenges. A major limitation is that spacing inherently requires more time, as increasing the intervals between sessions extends the overall study duration. This presents challenges for students with tight schedules, impending deadlines, or competing priorities, necessitating strategies to maximize learning outcomes within constrained time frames. Extant research on the spacing and testing effects controls the number of repetitions across conditions and varies total study time (for reviews of the spacing effect, see Cepeda et al., 2006; Donovan & Radosevich, 1999; Janiszewski et al., 2003; Maddox, 2016; for reviews of the testing effect, see Adesope et al., 2017; Karpicke & Aue, 2015; Rowland, 2014), leaving a key gap: how should learners allocate study time effectively in time-constrained settings? Specifically, should they prioritize fewer, widely spaced repetitions or more frequent repetitions with shorter intervals?

To the best of our knowledge, this trade-off has never been investigated before and is key to implementing spaced retrieval practices in meaningful environments. Additionally, item difficulty may influence the balance between repetition and spacing, as harder items often require more retrieval opportunities to achieve stable retention (Hoyer et al., 2003; Onyper et al., 2008). This study addresses the question by investigating the effects of spaced retrieval practice within a constrained time frame, with a particular focus on how item difficulty modulates the relationship between spacing intervals and practice frequency. The findings contribute to the development of more effective spaced retrieval practice interventions in meaningful learning and advance theoretical understanding of how these principles operate under practical constraints.

The Benefit of Increasing Spacing Intervals

The benefits of spaced retrieval practice can be understood through the lens of memory strengthening and retrieval dynamics. The study-phase retrieval theory posits that spacing is effective when the subsequent presentation brings to mind and is connected with the prior presentations (Appleton-Knapp et al., 2005; Thios & D'Agnostino, 1976; see also reminding theory, Benjamin & Tullis, 2010). According to the idea of desirable difficulty (Bjork, 1994;

In D. Barner, N.R. Bramley, A. Ruggeri and C.M. Walker (Eds.), *Proceedings of the 47th Annual Conference of the Cognitive Science Society*

Bjork & Bjork, 1992, 2011), spacing intervals between retrieval opportunities allow time for some forgetting to occur, which makes subsequent retrievals more effortful, and as a result, more impactful in strengthening memory traces (Latimier et al., 2021; Barzagar Nazari & Ebersbach, 2019).

However, this process depends on retrieval success: if intervals are too long, creating an "undesirable" level of difficulty, learners are less likely to recall the information, limiting the reinforcing effect of retrieval (Bjork & Bjork, 2020; Maddox et al., 2018; Pavlik & Anderson, 2005; Yan et al., 2024). For example, Verkoeijen et al. (2005) found that the benefits of increasing spacing intervals follow an inverted U-shaped pattern: free recall performance improved as spacing intervals increased to an optimal point, after which performance declined with further increases in spacing. Therefore, to maximize learning, spacing intervals must remain within an optimal range that balances effortful retrieval with achievable success.

The Benefit of Increasing the Amount of Practice

Even though retrieval practice is widely recognized as a powerful tool for enhancing memory, increasing the amount of practice does not always yield proportional gains. The relationship between practice amount and learning follows a nonlinear trajectory: improvement is rapid initially but slows as proficiency increases (e.g., Thorndike, 1913; Heathcote et al., 2000). This pattern, often described as the "power law of practice," reflects how performance systematically improves with practice (e.g., Ritter & Schooler, 2001; Stafford & Dewar, 2014). The ACT–R memory model developed by Anderson and Lebiere (1998) provides a theoretical framework for this phenomenon, positing that each practice session strengthens memory through increments that decay over time, cumulatively determining memory strength.

Laboratory studies also suggest that the "more-is-better" principle for practice is valid only to a certain extent (Lyle et al., 2020). Researchers have explored whether increasing the criterion level—the number of times an item must be correctly retrieved—leads to better memory performance (e.g., Karpicke & Roediger, 2008; Pyc & Rawson, 2009; Rawson & Dunlosky, 2011; Vaughn & Rawson, 2011). For example, Pyc and Rawson (2009) found that three retrievals improved performance compared to one retrieval when learning Swahili-English word pairs. However, as the number of retrievals increased from 5 to 10, the performance gains diminished on retention tests conducted 25 minutes later.

The challenge, therefore, lies in calibrating the balance between retrieval frequency and spacing intervals. While frequent repetition ensures that learners encounter the material often enough to stabilize memory, excessive repetition with short intervals can result in diminished returns. On the other hand, spacing that is too wide increases the risk of forgetting, making retrieval more likely to fail and reducing its reinforcing effect.

Item difficulty further complicates this balance. Research shows that harder items are typically associated with lower initial success rates, requiring more retrieval opportunities to achieve stable retention (Warr, 1963, 1964). These findings suggest that for challenging materials, increasing the frequency of retrieval opportunities with shorter spacing intervals may help learners consolidate memory and reduce the likelihood of retrieval failure. Conversely, for easier materials, incorporating wider intervals between retrievals can enhance retention without unnecessary repetition. By tailoring learning schedules to the characteristics of the material and the learner, it is possible to maximize memory strengthening while minimizing wasted time.

The Present Study

Across two experiments, while maintaining the total study time constant, we varied spacing levels and the number of repetitions to determine their combined effects on students' learning of foreign language word pairs. Item difficulty was manipulated with half the items categorized as difficult and half as easy. To ensure the total learning duration remained consistent across conditions, filler items were included to maintain control over spaced intervals. We hypothesized a significant interaction between repetition/spacing and item difficulty on learning outcomes. Specifically, we predicted that harder items would benefit more from frequent repetitions with shorter spacing intervals, as these provide sufficient retrieval opportunities to stabilize memory traces. Easier items, by contrast, were expected to benefit from fewer repetitions paired with longer spacing, as their stronger initial encoding reduces the likelihood of retrieval failure.

In Experiment 1, medium-difficulty foreign language pairs served as filler items. To eliminate the potential confounding effects of using similar learning content as fillers, Experiment 2 utilized math problems as filler items.

Experiment 1

Method

Experiment 1 was not pre-registered, but Experiment 2 was. The preregistration for Experiment 2, along with the raw data, R code, and supplemental materials, can be accessed on OSF (https://osf.io/dgsr8/). The studies were approved by the Athabasca University Research Ethics Board and Carnegie Mellon University's Institutional Review Board.

Participants A priori power analysis using G*Power 3.1 (Faul et al., 2009) estimated a required sample size of 199 for a small-to-medium effect (f = 0.2, α = .05, power = 80%) across 8 groups (4 repetition/spacing × 2 item difficulty). Since G*Power tends to underestimate sample sizes for interactions, we multiplied by 4, yielding 796, rounded up to 800. Participants were recruited through Prolific.com, and a total of 793 participants (456 female, 336 male, and 1 undisclosed) with a mean age of 28.78 years (SD = 4.43) completed the experiment. All were from English-speaking countries (Canada, the UK, and the US) and received $3.85 for ~20 minutes of participation. They were randomly assigned to one of the repetition/spacing conditions. Eleven participants were excluded due to prolonged inactivity (no-response trials exceeding 3 SDs above the sample's mean maximum no-response trials, excluding trivia questions). This resulted in a final sample of 782 participants: 194 (2 repetitions, 36-trial spacing), 199 (3 repetitions, 24-trial spacing), 199 (4 repetitions, 18-trial spacing), and 190 (6 repetitions, 12-trial spacing). For simplicity, these conditions were represented by the number of repetitions in the following sections.

Stimuli We used Indonesian-English word pairs from Kornell & Son (2009), selected based on existing difficulty norms. Twelve pairs were used as target items: six difficult items (0% accuracy after 5-8 trials) and six easy items (100% accuracy after 6-8 trials). An additional 24 medium-difficulty pairs (M = 42%, SD = 6.6% after 5–9 trials) were used as fillers, with the number varying by condition (see red blocks in Figure 1). When fewer than 24 fillers were needed, a random subset was selected from the full set. Eight trivia questions from Tauber et al. (2013) were presented between the learning phase and final test (see Supplementary Material).

Condition	72 training trials							
2 Repetitions	12 target	24	12 target	24				
3 Repetitions	12 target	12	12 target	12	12 target	12		
4 Repetitions	12 target	6	12 target	6	12 target	6	12 target	6
6 Repetitions	12 target	12 target	12 target	12 target	12 target	12 target		

Figure 1: Number of target and filler items in the learning session.

Experiment Design The experiment was a 2 (item difficulty) x 4 (repetition/spacing) mixed design, with spacing/repetition manipulated between participants with four levels (2 repetitions, 3 repetitions, 4 repetitions, and 6 repetitions) and item difficulty manipulated within participants with two levels (easy vs. hard).

We manipulated the trade-off between spacing and repetition across 72 learning trials (see Figure 1): items with larger spacing intervals were repeated fewer times, while items with smaller spacing intervals were repeated more frequently. Specifically, items repeated twice, three times, four times, and six times had average spacing intervals of 36, 24, 18, and 12 trials, respectively.

Procedure Participants were randomly assigned to four repetition/spacing conditions. All of them completed a learning session with 72 trials. In each trial, they were shown an Indonesian word and a space to enter the corresponding English translation. They were given 10 seconds to input the answer. Feedback displaying the correct answer was presented for 2 seconds.

After the learning session, participants completed a set of 8 trivia questions within 2 minutes (15 seconds per question). Following the trivia, participants completed a final test consisting of 36 trials: 12 target words (random order) followed by 24 fillers (random order). Participants had 10 seconds to input their answers, and no feedback was provided.

Results

We assessed response accuracy using Hypertext Preprocessor (PHP)'s *similar_text* function, which calculates the percentage of shared substrings between responses and correct answers. A similarity score of ⩾0.85 was considered correct. This threshold accounted for minor typos and regional spelling differences (e.g., theater vs. theatre).

Learning Session To examine the effects of repetition/spacing and item difficulty on participants' accuracy during the learning session, we conducted a 2 item difficulty × 4 repetition/spacing mixed-effects ANOVA using the *mixed* function in the *afex* package (Singmann et al., 2023) in R (R Core Team, 2024). Participant ID was included as a random effect, and repetition/spacing was treated as a continuous variable based on the repetition values. Due to the lack of consensus on standard effect sizes for mixed models, we report semi-partial R² for each fixed effect using *r2beta* from the *r2glmm* package, reflecting the unique variance explained while accounting for random effects. The results revealed a significant main effect of repetition/spacing, $F(1,780) = 964.46$, $p < .001$, $R_{sp}^2 = .525$, indicating that the increase of repetition with decreasing spacing intervals led to significantly increased accuracy during the learning session. For example, as shown in Figure 2, the 6-repetition condition (red line) consistently achieved higher accuracy than conditions with fewer repetitions. The main effect of word difficulty was also significant, $F(1,780) = 2450.24$, $p < .001$, $R_{sp}^2 = .392$, with better performance on the easy items (*M* = .48, *SE* = .005, 95% CI [.472, .494]) compared to the hard items (*M* = .24, *SE* = .005, 95% CI [.227, .248]). Furthermore, the interaction between repetition/spacing and difficulty was significant, $F(1,780) = 11.45$, $p < .001$, $R_{sp}^2 = .003$, indicating that the impact of repetition/spacing on accuracy differed depending on item difficulty. This interaction is depicted in Figure 2, which shows participants' accuracy on each trial during the learning session (e.g., the 6-repetition condition includes accuracy from the 1st to the 6th repetition). For easy items, performance improved quickly across the first few repetitions, with diminishing returns after the fourth repetition. In contrast, for hard items, accuracy steadily improved across all repetitions, with gains persisting until the final repetition.

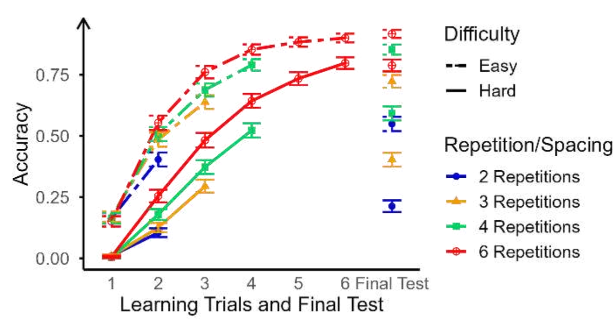

Figure 2: Participants' average accuracy in Experiment 1
Note. Error bars show ± 2SE.

To further investigate the interaction between item difficulty and repetition condition, we conducted a simple slopes analysis to examine the effect of repetition/spacing on learning performance for each difficulty level. The estimated slope for easy items was b = 0.098, 95% CI [0.090, 0.105], and for the hard items, b = 0.109, 95% CI [0.102, 0.116]. A comparison of the slopes revealed a significant difference, $t(780) = -3.38$, $p < .001$, with the slope for the hard items being steeper than for the easy items by $\Delta b = -0.011$, SE = .003, which suggests that during learning, difficult items benefit more from additional repetitions and less spacing.

Final Test We also conducted the two-way mixed ANOVA to examine the effects of repetition/spacing and word difficulty on accuracy for the final test. Similar to what we found in the learning session, the results of the final test showed that there was a significant main effect of repetition/spacing, $F(1, 780) = 504.81$, $p < .001$, $R_{sp}^2 = .328$. The main effect of word difficulty was also significant, $F(1, 780) = 941.75$, $p < .001$, $R_{sp}^2 = .227$, suggesting that easy items (M = .76, SE = .009, 95% CI [.743, .777]) were significantly higher than hard items (M = .50, SE = .009, 95% CI [.480, .514]). Additionally, there was a significant interaction between repetition/spacing and word difficulty, $F(1, 780) = 85.64$, $p < .001$, $R_{sp}^2 = .026$. A simple slopes analysis revealed that hard items (b = 0.143, SE = 0.006, 95% CI [0.131,0.154]) showed a steeper performance increase than easy items (b = 0.089, SE = 0.006, 95% CI [0.077,0.100]) as repetitions increased, $t(780) = 9.25$, $p < .001$, $\Delta b = 0.054$, SE = 0.006.

Discussion

The findings of Experiment 1 showed that, within a fixed amount of time, both easy and hard items benefit from an increased number of repetitions with smaller spacing intervals for both the learning session and the final test. However, hard items exhibited a greater benefit from additional repetitions compared to easy items. These results support our hypothesis for the interaction between the repetition/spacing condition and item difficulty: difficult items would benefit more from a greater number of repetitions with smaller spacing intervals than easy items. This finding is consistent with previous research indicating that more difficult items, often associated with lower initial success rates, require more practice and frequent engagement to achieve the same level of mastery as easier items (Warr, 1963, 1964).

Interestingly, easy items also demonstrated benefits from more repetitions. This outcome can be attributed to the fact that, although the items classified as "easy" in our study were designed to be less challenging compared to the "difficult" items, they were still novel to participants, who had no prior knowledge of Indonesian. Unlike studies where participants engage in spaced practice for material they have already learned to strengthen long-term memory retention (e.g., Cepeda et al., 2008), our study involved initial exposure to entirely new material. As a result, participants benefited from increased repetitions. Furthermore, we found that the improvement rate for easy items slowed with additional repetitions, a pattern aligning with prior findings that diminishing returns occur as practice progresses (e.g., Pyc & Rawson, 2009).

A potential confounding variable in Experiment 1 was the varying number of filler items, which may have influenced accuracy. Participants, unaware of the distinction between target and filler words, practiced with different numbers of words across conditions, potentially affecting performance. For instance, participants in the 2-repetition condition were exposed to 36 unique items (12 target words and 24 filler items), while those in the 6-repetition condition encountered only 12 unique items (no filler items). Consequently, the cognitive load differed between conditions: participants in lower-repetition conditions may have performed worse due to the greater number of items to learn, whereas those in higher-repetition conditions may have performed better with fewer items to learn. To address this issue, we conducted Experiment 2, replacing the filler items with math problems to eliminate the additional memory load.

Experiment 2

Experiment 2 was identical to Experiment 1, except that the filler items were replaced with math problems. Participants were informed that their task was to learn the Indonesian words, and they would be tested only on those later. All analyses reported below, unless otherwise stated, follow the pre-registration.

Method

A total of 796 participants from Prolific.com completed the experiment. We used the same cutoff threshold as Experiment 1. Four participants were excluded due to being inactive (no response) for a longer period of time. Additionally, two participants were excluded due to missing trials. Finally, 792 participants were included in the analysis. The number of participants in each condition was as follows: 200 (2 repetitions), 201(3 repetitions), 188 (4 repetitions), and 201 (6 repetitions).

Stimuli The stimuli were the same as in Experiment 1, consisting of 12 Indonesian-English word pairs used as target items. The key difference was the use of 24 math problems as fillers, which are simple calculations (see Supplementary Material).

Results

Learning Session We conducted the two-way mixed ANOVA as we did in Experiment 1 to examine the effects of repetition/spacing and word difficulty on accuracy for the learning session. This deviated from the pre-registration, which specified treating repetition/spacing as a categorical variable. The reason was that we aimed to examine potential trends in the data rather than focusing on how specific repetition values affect word learning. The results showed a significant main effect of repetition/spacing, $F(1, 788) = 714.07$, $p < .001$, $R_{sp}^2 = .423$, indicating that repetition/spacing had a significant effect on accuracy during the learning session. The main effect of word difficulty was also significant, $F(1, 788) = 2720.82$, $p < .001$, $R_{sp}^2 = .394$, suggesting that easy items ($M = .52$, $SE = .005$, 95% CI [.507, .528]) were significantly higher than hard items ($M = .27$, $SE = .005$, 95% CI [.260, .282]). The interaction between repetition/spacing and word difficulty was also significant, though the effect size was small, $F(1, 788) = 8.22$, $p = .004$, $R_{sp}^2 = .002$**.** Figure 3 depicts participants' accuracy, including the learning session and final test. Overall, the general pattern remained consistent with Experiment 1. However, a key difference was that participants' accuracy across repetition/spacing conditions in Experiment 2 showed greater overlap within the same number of repetitions for both easy and hard items. This suggests that math filler items did not affect participants' performance in Experiment 2, leading to more comparable accuracy levels across conditions.

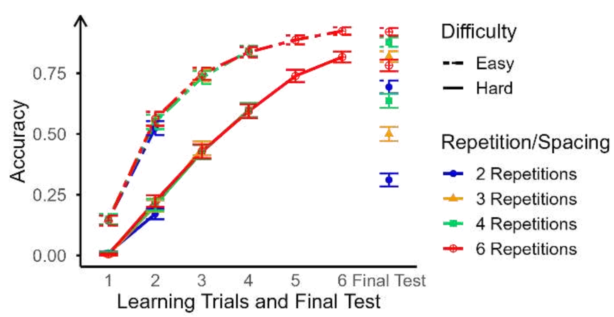

Figure 3: Participants' average accuracy in Experiment 2
Note. Error bars show ± 2SE.

A simple slopes analysis was conducted to examine the effect of repetition/spacing on learning performance for each difficulty level. This also deviated from the pre-registration, which indicated that multiple comparisons between different conditions would be conducted using the BH correction. The change was made because treating repetition/spacing as a continuous variable eliminates the need for multiple comparisons. Even though the slope of the hard items ($b = 0.092$, $SE = 0.004$, 95% CI [0.085,0.100]) was still significantly higher than the easy items ($b = 0.083$, $SE = 0.004$, 95% CI [0.076,0.090]), $t(788) = -2.87$, $p = .004$, the difference was small ($\Delta b = -0.009$, $SE = 0.003$).

Final Test We conducted a two-way mixed ANOVA to examine the effects of repetition/spacing and word difficulty on accuracy for the final test. Repetition/spacing was treated as a continuous variable, deviating from the pre-registration for the reasons mentioned earlier. The results still showed a significant main effect of repetition/spacing, $F(1, 788) = 278.16$, $p < .001$, $R_{sp}^2 = .208$. The main effect of word difficulty was also significant, $F(1, 788) = 953.27$, $p < .001$, $R_{sp}^2 = .237$, suggesting that easy items ($M = .83$, $SE = .009$, 95% CI [.810, .844]) were significantly higher than hard items ($M = .56$, $SE = .009$, 95% CI [.539, .573]). Additionally, there was a significant interaction between repetition/spacing and word difficulty, $F(1, 788) = 108.40$, $p < .001$, $R_{sp}^2 = .034$. The simple slopes analysis was a deviation from pre-registration for the same reasons previously discussed. The estimated slope for the easy items ($b = 0.053$, $SE = 0.006$, 95% CI [0.042,0.064]) was still significantly smaller than the hard items ($b = 0.114$, $SE = 0.006$, 95% CI [0.103,0.126]), $t(788) = -10.41$, $p < .001$, $\Delta b = -0.061$, $SE = 0.006$.

Discussion

In Experiment 2, math problems were used as filler items to control the potential confounding effects of the number of practiced items across different conditions. The findings were consistent with those of Experiment 1, suggesting that the observed differences between repetition/spacing conditions were primarily driven by the number of repetitions and spacing intervals, rather than the number of unique items practiced. Both easy and hard items benefited from an increased number of repetitions with smaller spacing intervals. However, hard items showed a greater benefit from additional repetitions compared to easy items, for which the rate of improvement slowed as repetitions increased.

General Discussion

Across two experiments, we consistently observed that increasing the number of repetitions rather than spacing intervals improved learning outcomes for both easy and hard items, particularly in time-constrained settings. The observed benefits of increased repetitions over larger spacing intervals suggest that, when exposed to novel learning content, learners need to revisit the material multiple times to develop proficiency. This finding aligns with the well-established fact that repetition enhances retention, first demonstrated in Ebbinghaus's experiments and subsequently validated by extensive research (Hintzman, 1976). It also supports the power law of practice, which states that increased practice during the initial stages of learning yields substantial benefits (e.g., Ritter & Schooler, 2001; Stafford & Dewar, 2014). Furthermore, our findings align with prior research showing that less well-learned material benefits more from shorter intervals between repetitions (Verkoeijen et al.,2005; Carpenter et al., 2022).

The retrieval-based theories (Appleton-Knapp et al., 2005; Thios & D'Agnostino, 1976; see also reminding theory, Benjamin & Tullis, 2010; Hintzman, 2011) provide a theoretical framework for understanding the potential mechanisms underlying the balance between repetition and spacing. According to these theories, the spacing of repeated presentations may benefit memory because later exposures can remind learners of earlier ones, thereby reinforcing memory traces. However, the effectiveness of this process likely depends on the timing and frequency of repetitions: items successfully retrieved or reminded during practice tend to be better remembered on final tests compared to those where retrieval fails (e.g., Jacoby, 1974; Maddox et al., 2018; Wahlheim et al., 2014). This suggests that insufficient repetition can prevent learners from retrieving previously encountered items, leading to weaker memory enhancement.

Similarly, excessively long delays between repetitions may lead to forgetting, preventing effective reminding of earlier encounters. Therefore, during the initial stages of learning, learners may require more frequent repetitions with shorter spacing intervals to ensure successful retrieval and consolidation. This approach minimizes the risk of forgetting while reinforcing memory traces through repeated engagement.

Another important finding across both experiments is the interaction between item difficulty and repetition/spacing. This interaction reveals that the effect of increasing repetitions while reducing spacing intervals is influenced by item difficulty. Harder items, associated with lower initial success rates, benefit more from increased repetitions and shorter spacing intervals to achieve mastery. In contrast, easier items showed diminishing returns with additional repetitions, consistent with the power law of practice (e.g., Lyle et al., 2020). These results align with prior research showing that harder items require more practice and frequent engagement to achieve the same level of mastery as easier items (Warr, 1963, 1964).

The stronger positive relationship between repetition/spacing and final test performance for harder items also underscores the need to consider item difficulty when optimizing spacing intervals. While the idea of desirable difficulty (Bjork, 1994; Bjork & Bjork, 1992, 2011) emphasizes the benefits of retrieval effort, it does not fully account for how item difficulty interacts with spacing and repetition. Our findings suggest that item difficulty significantly influences these dynamics. For harder items with weaker initial encoding, more retrieval opportunities with smaller spacing intervals increase the likelihood of successful retrieval, reinforcing memory traces more effectively. In contrast, for easier items with stronger initial encoding, larger spacing intervals with fewer repetitions can make retrieval effortful yet successful, enhancing memory retention without redundancy. As practice progresses, difficult items may become easier, allowing spacing intervals to be gradually increased to maintain optimal retrieval effort.

Practical Implications

Taken together, our findings not only advance the theoretical understanding of spaced retrieval practice but also offer practical guidance for designing effective learning schedules. In time-constrained scenarios, prioritizing frequent repetitions over extended spacing intervals appears more beneficial initially, particularly for challenging material. For easier content, learners may benefit from frequent repetition in the beginning, but once they have had sufficient exposure, they may benefit more from wider spacing intervals and fewer repetitions. Therefore, designing effective spaced retrieval practice schedules within a limited time requires balancing spacing intervals, repetition frequency, and item difficulty. Achieving this balance can help ensure that learners experience desirable difficulties that promote effortful, yet successful, retrieval processes and avoid conditions where excessive difficulty or redundancy hinders learning.

These findings also have significant implications for the design of adaptive learning platforms, educational tools, and instructional strategies. Learning platforms, such as flashcard apps or language-learning programs, can leverage these results by adjusting spacing intervals and repetition frequency based on item difficulty, learner performance, and learners' time availability and deadlines (e.g., exam date). For example, harder items could receive more repetitions with shorter intervals or spaced practice to stabilize memory, while easier items could shift to longer spacing as mastery improves. Similarly, educational curricula and training programs could prioritize frequent repetition of challenging material during initial learning stages, gradually increasing spacing for easier content as learners progress. This approach can also inform professional training modules and standardized exam preparation, where study schedules could emphasize frequent engagement with complex content while balancing the time constraints learners face. By tailoring study schedules to material difficulty and retention goals, these strategies can help optimize learning outcomes across a variety of real-world contexts.

Limitations and Future Directions

One limitation of our study is that we examined participants' final test performance at an immediate test. Previous research suggests that the benefits of spaced practice are more pronounced for long-term retention (e.g., Greving & Richter, 2019; Rohrer & Taylor, 2006). Future research should extend this work by examining how retention intervals influence the repetition-spacing trade-off, building on existing findings on item difficulty. Investigating these interactions can provide a more comprehensive understanding of how to optimize spaced retrieval practice for real-world learning contexts and goals.

Conclusion

The limiting factor learners often face is time—whether it's about the amount of time they have to study before an exam, the free time they have available to engage in a hobby, the number of weeks they have before they take a trip where they need to use newly learned foreign vocabulary, or simply the amount of time they are willing to take away from other important things in one's life. We know from cognitive literature that repetition and spacing are both important, but given a fixed time frame, learners must decide how to balance one against the other. Increased repetitions come with reduced spacing and vice versa. This study highlights the critical trade-off and shows that optimizing the balance is also influenced by item difficulty: the harder the items, the more important it is to favor repetitions over spacing. Our results underscore the significance of tailoring learning strategies to both the difficulty of the material and the constraints of the learning context, providing valuable insights for building effective learning systems and educational practices.

Acknowledgments

This research was supported by funding from the Social Sciences and Humanities Research Council of Canada (SSHRC) through an Insight Grant (# 435-2021-0426, PI: Faria Sana), NSF Grant (#2238567, PI: Veronica Yan), and NSF DRL Grant (#2301130, PI: Paulo F. Carvalho).

References

Adesope, O. O., Trevisan, D. A., & Sundararajan, N. (2017). Rethinking the use of tests: A meta-analysis of practice testing. *Review of educational research*, *87*(3), 659-701. https://doi.org/10.3102/0034654316689306

Anderson, J. R., & Lebiere, C. (1998). *Atomic components of thought*. Hillsdale, NJ: Erlbaum.

Appleton-Knapp, S. L., Bjork, R. A., & Wickens, T. D. (2005). Examining the spacing effect in advertising: Encoding variability, retrieval processes, and their interaction. *Journal of Consumer Research*, *32*(2), 266-276. https://doi.org/10.1086/432236

Barzagar Nazari, K., & Ebersbach, M. (2019). Distributing mathematical practice of third and seventh graders: Applicability of the spacing effect in the classroom. *Applied Cognitive Psychology*, *33*(2), 288-298. https://doi.org/10.1002/acp.3485

Benjamin, A. S., & Tullis, J. G. (2010). What makes distributed practice effective? *Cognitive Psychology*, *61*,228–247. https://doi.org/10. 1016/j.cogpsych.2010.05.004

Bjork, R. A., & Bjork, E. L. (1992). A new theory of disuse and an old theory of stimulus fluctuation. In A. F. Healy, S. M. Kosslyn, & R. M. Shiffrin (Eds.), *Essays in honor of William K. Estes, Vol. 1. From learning theory to connectionist theory; Vol. 2. From learning processes to cognitive processes* (pp. 35–67). Lawrence Erlbaum Associates, Inc.

Bjork, R. A. (1994). Memory and metamemory considerations in the training of human beings. In J. Metcalfe & A. Shimamura (Eds.), Metacognition: Knowing about knowing (pp. 185–205). MIT Press.

Bjork, E. L., & Bjork, R. A. (2011). Making things hard on yourself, but in a good way: Creating desirable difficulties to enhance learning. In M. A. Gernsbacher, R. W. Pew, L. M. Hough, & J. R. Pomerantz (Eds.), *Psychology and the real world: Essays illustrating fundamental contributions to society* (pp. 56-64). New York: Worth Publishers.

Bjork, R. A., & Bjork, E. L. (2020). Desirable difficulties in theory and practice. *Journal of Applied Research in Memory and Cognition,* *9*(4), 475–479. https://doi.org/10.1016/j.jarmac.2020.09.003

Carpenter, S. K., Pan, S. C., & Butler, A. C. (2022). The science of effective learning with spacing and retrieval practice. *Nature Reviews Psychology*, *1*(9), 496-511. https://doi.org/10.1038/s44159-022-00089-1

Cepeda, N. J., Pashler, H., Vul, E., Wixted, J. T., & Rohrer, D. (2006). Distributed practice in verbal recall tasks: A review and quantitative synthesis. *Psychological Bulletin*, 132(3), 354.

Cepeda, N. J., Vul, E., Rohrer, D., Wixted, J. T., & Pashler, H. (2008). Spacing effects in learning: A temporal ridgeline of optimal retention. *Psychological science*, *19*(11), 1095-1102. https://doi.org/10.1111/j.1467-9280.2008.02209.x

Donovan, J. J., & Radosevich, D. J. (1999). A meta-analytic review of the distribution of practice effect: Now you see it, now you don't. *Journal of Applied Psychology*, *84*(5), 795. https://doi.org/10.1037/0021-9010.84.5.795

Dunlosky, J., Rawson, K. A., Marsh, E. J., Nathan, M. J., & Willingham, D. T. (2013). Improving students' learning with effective learning techniques: Promising directions from cognitive and educational psychology. *Psychological Science in the Public interest*, *14*(1), 4-58. https://doi.org/10.1177/1529100612453266

Faul, F., Erdfelder, E., Buchner, A., & Lang, A. G. (2009). Statistical power analyses using G* Power 3.1: Tests for correlation and regression analyses. *Behavior research methods*, *41*(4), 1149-1160. https://doi.org/10.3758/BRM.41.4.1149

Greving, C. E., & Richter, T. (2019). Distributed learning in the classroom: Effects of rereading schedules depend on time of test. *Frontiers in Psychology*, *9*, 2517. https://doi.org/10.3389/fpsyg.2018.02517

Heathcote, A., Brown, S., & Mewhort, D. J. (2000). The power law repealed: The case for an exponential law of practice. *Psychonomic bulletin & review*, *7*(2), 185-207. https://doi.org/10.3758/BF03212979

Hintzman, D. L. (1976). Repetition and memory. *Psychology of learning and motivation*, *10*, 47-91. https://doi.org/10.1016/S0079-7421(08)60464-8

Hintzman, D. L. (2011). Research strategy in the study of memory: Fads, fallacies, and the search for the "coordinates of truth". *Perspectives on Psychological Science*, *6*(3), 253-271. https://doi.org/10.1177/1745691611406924

Hoyer, W., Cerella, J., & Onyper, S. (2003). Item learning in cognitive skill training: Effects of item difficulty. *Memory & Cognition*, *31*, 1260-1270. https://doi.org/10.3758/BF03195809.

Jacoby, L. L. (1974). The role of mental contiguity in memory: Registration and retrieval effects. *Journal of Verbal Learning and Verbal Behavior*, *13*, 483–496. https://doi.org/10.1016/S00225371(74)80001-0

Janiszewski, C., Noel, H., & Sawyer, A. G. (2003). A meta-analysis of the spacing effect in verbal learning: Implications for research on advertising repetition and consumer memory. *Journal of Consumer Research*, *30*(1), 138-149. https://doi.org/10.1086/374692

Karpicke, J. D., & Aue, W. R. (2015). The testing effect is alive and well with complex materials. *Educational Psychology Review*, *27*, 317-326. https://doi.org/10.1007/s10648-015-9309-3

Karpicke, J. D., & Roediger III, H. L. (2008). The critical importance of retrieval for learning. *science*, *319*(5865), 966-968. https://doi.org/10.1126/science.1152408

Kornell, N., & Son, L. K. (2009). Learners' choices and beliefs about self-testing. *Memory*, *17*(5), 493-501. https://doi.org/10.1080/09658210902832915

Latimier, A., Peyre, H., & Ramus, F. (2021). A Meta-Analytic Review of the Benefit of Spacing out Retrieval Practice Episodes on Retention. *Educational Psychology Review*, *33*(3), 959–987. https://doi.org/10.1007/s10648-020-09572-8

Lyle, K. B., Bego, C. R., Hopkins, R. F., Hieb, J. L., & Ralston, P. A. (2020). How the amount and spacing of retrieval practice affect the short-and long-term retention of mathematics knowledge. *Educational Psychology Review*, *32*, 277-295. https://doi.org/10.1007/s10648-019-09489-x

Maddox, G. B. (2016). Understanding the underlying mechanism of the spacing effect in verbal learning: A case for encoding variability and study-phase retrieval. *Journal of Cognitive Psychology*, *28*(6), 684-706. https://doi.org/10.1080/20445911.2016.1181637

Maddox, G. B., Pyc, M. A., Kauffman, Z. S., Gatewood, J. D., & Schonhoff, A. M. (2018). Examining the contributions of desirable difficulty and reminding to the spacing effect. *Memory & Cognition*, *46*, 1376-1388. https://doi.org/10.3758/s13421-018-0843-3

Onyper, S., Hoyer, W., & Cerella, J. (2008). Effects of Item Difficulty on the Retrieval of Solutions during Cognitive Skill Acquisition: Age Differences. *Aging, Neuropsychology, and Cognition*, *15*, 358 - 383. https://doi.org/10.1080/13825580701879907.

Pashler, H., Rohrer, D., Cepeda, N. J., & Carpenter, S. K. (2007). Enhancing learning and retarding forgetting: Choices and consequences. *Psychonomic Bulletin & Review*, *14*(2), 187-193.
https://doi.org/10.3758/BF03194050

Pavlik Jr, P. I., & Anderson, J. R. (2005). Practice and forgetting effects on vocabulary memory: An activation-based model of the spacing effect. *Cognitive science*, *29*(4), 559-586. https://doi.org/10.1207/s15516709cog0000_14

Pyc, M. A., & Rawson, K. A. (2009). Testing the retrieval effort hypothesis: Does greater difficulty correctly recalling information lead to higher levels of memory?. *Journal of Memory and Language*, *60*(4), 437-447. https://doi.org/10.1016/j.jml.2009.01.004

Rawson, K. A., & Dunlosky, J. (2011). Optimizing schedules of retrieval practice for durable and efficient learning: How much is enough? *Journal of Experimental Psychology: General,* *140*(3), 283–302. https://doi.org/10.1037/a0023956

R Core Team (2024). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. https://www.R-project.org/

Ritter, F. E., & Schooler, L. J. (2001). The learning curve. *International encyclopedia of the social and behavioral sciences*, *13*, 8602-8605.

Rohrer, D., & Taylor, K. (2006). The effects of overlearning and distributed practise on the retention of mathematics knowledge. *Applied Cognitive Psychology: The Official Journal of the Society for Applied Research in Memory and Cognition*, *20*(9), 1209-1224. https://doi.org/10.1002/acp.1266

Rowland, C. A. (2014). The effect of testing versus restudy on retention: A meta-analytic review of the testing effect. *Psychological Bulletin,* *140*(6), 1432–1463. https://doi.org/10.1037/a0037559

Singmann, H., Bolker, B., Westfall, J., Aust, F., & Ben-Shachar, M. S. (2023). *afex: Analysis of factorial experiments* (Version 1.3-0) [R package]. CRAN. https://CRAN.R-project.org/package=afex

Stafford, T., & Dewar, M. (2014). Tracing the trajectory of skill learning with a very large sample of online game players. *Psychological science*, *25*(2), 511-518. https://doi.org/10.1177/0956797613511466

Tauber, S. K., Dunlosky, J., Rawson, K. A., Rhodes, M. G., & Sitzman, D. M. (2013). General knowledge norms: Updated and expanded from the Nelson and Narens (1980) norms. *Behavior research methods*, *45*, 1115-1143. https://doi.org/10.3758/s13428-012-0307-9

Thios, S. J., & D'Agostino, P. R. (1976). Effects of repetition as a function of study-phase retrieval. *Journal of Verbal Learning and Verbal Behavior*, *15*(5), 529-536. https://doi.org/10.1016/0022-5371(76)90047-5

Thorndike, E. L. (1913). *Educational psychology: The psychology of learning* (Vol. 2). New York: Teachers College Press.

Vaughn, K. E., & Rawson, K. A. (2011). Diagnosing criterion-level effects on memory: What aspects of memory are enhanced by repeated retrieval?. *Psychological Science*, *22*(9), 1127-1131. https://doi.org/10.1177/0956797611417724

Verkoeijen, P. P., Rikers, R. M., & Schmidt, H. G. (2005). Limitations to the spacing effect: Demonstration of an inverted u-shaped relationship between interrepetition spacing and free recall. *Experimental Psychology*, *52*(4), 257-263. https://doi.org/10.1027/1618-3169.52.4.257

Wahlheim, C. N., Maddox, G. B., & Jacoby, L. L. (2014). The role of reminding in the effects of spaced repetitions on cued recall: Sufficient but not necessary. *Journal of Experimental Psychology: Learning, Memory, and Cognition,* *40*(1), 94–105. https://doi.org/10.1037/a0034055

Warr, P. B. (1963). The effect of repeated presentation of items on paired associate learning. *Quarterly Journal of Experimental Psychology*, *15*(4), 262-272. https://doi.org/10.1080/17470216308416З

Warr, P. (1964). Repetition and Task Difficulty. Nature, 202, 422-422. https://doi.org/10.1038/202422A0.

Yan, V. X., Sana, F., & Carvalho, P. F. (2024). No Simple Solutions to Complex Problems: Cognitive Science Principles Can Guide but Not Prescribe Educational Decisions. *Policy Insights from the Behavioral and Brain Sciences*, *11*(1), 59-66. https://doi.org/10.1177/23727322231218906