Distributed Practice and Time Pressure Interact to Affect Learning and Retention of Arithmetic Facts

Olivia R. Jazbutis [1], Melody Wiseheart [2], Gabriel A. Radvansky [1], Nicole M. McNeil [1]

[1] *Department of Psychology, University of Notre Dame, Notre Dame, IN, USA.* [2] *Department of Psychology and LaMarsh Centre for Child and Youth Research, York University, Toronto, ON, Canada.*

Corresponding Author: Nicole M. McNeil, Department of Psychology, University of Notre Dame, 390 Corbett Family Hall, Notre Dame, IN 46556, USA. E-mail: nmcneil@nd.edu

Abstract

Arithmetic is commonly taught through timed practice and drill, yet little research exists to guide optimal practice structure. This study investigated the effects of distributed practice and time pressure on the acquisition and retention of arithmetic facts. Following a pretest, adult participants (n = 211) were randomly assigned to learn unfamiliar times tables (17 and 19) in one of ten conditions in a 5 (spacing: daily, every other day, weekly, every 10 days, every other week) x 2 (time pressure: timed or untimed) factorial design. After the learning phase, retention tests were given to measure both accuracy and response time immediately, after a ten-day delay, and at the end of semester. Time pressure during learning elevated participants' perceived stress. It also led to faster response times during testing when learning was spaced daily and every other day, but slower response times for all other spacings. These patterns were reversed in the absence of time pressure during learning. While timed and untimed practice during learning led to similar forgetting of practiced facts over time, untimed practice allowed participants to gradually improve on unpracticed facts and conceptually related facts across test phases. Ultimately, distributed practice and time pressure may interact in complex ways to affect the learning and retention of arithmetic facts, and the effects shown in previous studies using verbal material (e.g., narrative texts, word lists) may not generalize to arithmetic.

Keywords

Fluent retrieval of arithmetic facts (e.g., 5 + 8 = 13; 6 x 7 = 42) supports higher-level mathematics proficiency (Price et al., 2013). This idea is based on the Decomposition Thesis, which suggests that complex cognitive skills, such as those needed for solving differential equations or linear algebra problems, can be broken down into smaller component skills like arithmetic facts, and mastery of these component skills can improve performance in the larger domain (Anderson, 2002). Developing automaticity with arithmetic facts can lead to working memory savings that can be applied to solving more complex mathematical problems (Arroyo et al., 2011; Tronsky & Royer, 2003). Parents and educators often use tools such as flashcards, computer games, or informal activities to assist students in developing automaticity with arithmetic (LeFevre et al., 2009). The *What Works Practice Guide for Assisting Students Struggling with Mathematics* (Fuchs et al., 2021) recommends timed facts practice as part of a multi-component intervention, but further research is needed to identify the most effective ways to structure arithmetic practice.

Arithmetic Facts and Memory

Arithmetic facts are stored in generalized semantic memory (Tulving, 1972). However, cognitive processes other than direct retrieval are involved, such as the back-up strategies of estimation, conceptual subitizing, and calculation, so procedural memories also likely contribute to arithmetic memory (Gilmore et al., 2017). The Redescription Hypothesis posits that arithmetic is encoded by the associative system, and successive encounters activate associative representations, linking declarative and procedural knowledge (Dixon & Bangert, 2005). How children develop arithmetic knowledge may affect their conceptual and procedural knowledge of mathematics (Gilmore, 2006). Additionally, recall of arithmetic facts is affected by the context in which learning occurs, as episodic memories aid recall in similar contexts (Herbert & Burt, 2004).

When learning arithmetic, the sequence and visual attributes of numbers are initially encoded into working memory. As knowledge is acquired, memories consolidate in the hippocampus (Squire et al., 2015). Hippocampal memories undergo continuous strengthening, weakening, and other modifications (Horner & Doeller, 2017) through a process known as long-term potentiation (Lynch et al., 1990). According to systems consolidation theory, hippocampal-dependent memories are used to train cortical network representations by integrating new memories with existing ones in the cortex (Dudai et al., 2015). Eventually, aspects of new memories may be stored in more stable, hippocampus-independent representations, resulting in declarative knowledge (Sweegers et al., 2014).

The process of memory consolidation occurs over time, but the time course of this hippocampal-neocortical dialogue remains unknown (Squire et al., 2015). Two temporal factors that may affect the learning and retention of arithmetic facts are the interstudy interval (the time between learning sessions) and the retention interval (the time between the most recent learning and test); these factors may vary with the type of material (e.g., narrative texts, word lists, arithmetic; Raghubar et al., 2010). Hence, examining arithmetic retrieval may provide insights into the temporal dimensions of memory in mathematical cognition and the learning mechanisms involved. Given that arithmetic proficiency is helpful for academic achievement and daily life, it is crucial to identify the most effective ways to structure arithmetic practice. The present study examined two factors: (a) distributed practice and (b) time pressure.

Distributed Practice

Ebbinghaus (1885/1964) introduced the concept of the *distributed practice effect*, which suggests that dividing practice into multiple learning sessions leads to better memory retention when compared to massed practice, where the material is condensed into a single session. Widely studied since then, distributed practice cultivates superior long-term retention in comparison to massed repetitions, producing an advantage known as the *spacing effect* (Kang et al., 2016).

Studies investigating the distributed practice effect have mainly focused on verbal learning tasks, such as paired-associates, sentences, or text passages. A limited number of studies have explored the impact of distributed practice on arithmetic learning, and these have shown that it increases accuracy (Rea & Modigliani, 1985) and decreases response time (Rickard et al., 2008). Response time effects can occur without any impact on accuracy (Kwon et al., 2015), which may be due to the multifaceted nature of distributed practice: it provides a mental break while also challenging the mind with subsequent encounters.

There are several theories of distributed practice. The *Consolidation* account suggests that studying and restudying over an extended time allows memories to consolidate (Cohen et al., 2005; Korman et al., 2003). The *Deficient-Processing* account suggests that distributed practice reduces mind wandering and allows more attention to be directed to any active effort to retrieve memories (Hintzman, 1974; Metcalfe & Xu, 2016). The *Contextual Variability* account suggests that variation in internal and external learning contexts opens the possibility of multiple retrieval pathways (Glenberg, 1976, 1979). The *Study-Phase Retrieval* account suggests that making retrieval harder (e.g., from weaker representations due to partial forgetting) strengthens representations during relearning (Toppino et al., 2018). Thus, subsequent study sessions are an opportunity for learners to make connections, reactivating and strengthening memories to build upon prior knowledge (Benjamin & Tullis, 2010). Regardless of which of these theoretical views is most accurate, distributed practice is an opportunity for memory consolidation. This transformation of new material into long-term memory makes distributed practice an efficient strategy to improve long-term retention.

Research on distributed practice also has been conducted in practical settings (e.g., Goossens et al., 2016), as investigating the spacing effect in educational contexts can lead to valuable recommendations for improving the quality of learning (Wiseheart et al., 2019). Some previous recommendations for incorporating distributed practice in the classroom include splitting up lessons, revisiting information, tailoring technology, and administering cumulative assessments (Carpenter & Agarwal, 2020). Additionally, retention tests can serve as distributed practice sessions that enhance long-term retention (Roediger & Karpicke, 2006). In fact, several academic curricula, including Saxon mathematics, now incorporate distributed practice by distributing and integrating concepts to provide multiple learning opportunities (Bullock et al., 2015).

To provide more precise recommendations about optimal spacing, it is necessary to directly examine the timing of memory consolidation and retroactive interference as well as how this fits with the storage of memories in the hippocampus and the transfer of knowledge to the cortex (Squire et al., 2015). Some evidence suggests that hippocampal memory loss may occur approximately one week after learning new material (Fisher & Radvansky, 2018), although this is not observed for all types of material, such as well-integrated narratives (Fisher & Radvansky, 2019). According to Radvansky et al. (2022), memory retention passes through several phases, including working memory (WM) for time periods less than one minute, early Long-Term Memory (e-LTM) for time periods from one minute to 12 hours, transitional Long-Term Memory (t-LTM) for time periods from 12 hours to around seven days, and Long-Lasting Memory (LLM) for periods longer than around one week. However, the ability to observe the consequences of these transitions can vary depending on the complexity of the material, how well it was learned, and its similarities to previously learned material.

The concept of multiple memory phases, regardless of the exact number of phases, is consistent with the two-factor model, which integrates encoding variability and study-phase retrieval mechanisms into a single model (Toppino & Bloom, 2002). Encoding variability suggests that each new encounter with information is encoded differently, leading to supplementary retrieval cues and additional memory traces. Study-phase retrieval suggests that the first encounter must be retrieved from long-term memory during the second encounter to enhance the memory trace. This theory explains distributed practice intervals and predicts an inverted-U-shaped relationship between learning opportunities and free recall (Verkoeijen et al., 2008). As the time between sessions increases, so does free recall, up to a certain point, after which performance begins to decrease. Therefore, study-phase retrieval suggests that successive encounters should not be distributed more than three and a half weeks apart to maintain the benefits of distributed practice (Verkoeijen et al., 2008).

When planning for optimal spacing between learning sessions, it is important to consider the non-linear relationship between the length of time between learning sessions and the length of time between the last learning session and the test (Cepeda et al., 2008). Forgetting may vary based on this ratio, suggesting that the length of the optimal spacing interval is not fixed. This differs from the suggestion by Fisher and Radvansky (2018) that hippocampal memory loss is always observed one week after learning new material, particularly for simple facts presented once. However, it is consistent with more recent suggestions by Fisher and Radvansky (2019, 2022) and Radvansky et al. (2022) that the consequences of hippocampal memory loss may not always be observable one week after first learning complex material presented multiple times.

It is also consistent with the Multiscale Context Model (MCM) of long-term memory, which predicts how a study schedule influences retention and accounts for declarative memory (Mozer et al., 2009). While the MCM and two-factor model both aim to improve students' math proficiency through optimal scheduling, the models differ in their considerations: specifically, the two-factor model considers the inter-study interval while the MCM considers the retention interval and can treat it as a random variable when the actual retention interval is unknown. Thus, by integrating both models or determining which is best suited to learning arithmetic facts, educators can better tailor their teaching strategies to improve students' mathematics proficiency.

It is important to note that both the MCM and two-factor model are based on verbal learning and may not fully apply to the mechanisms involved in learning arithmetic facts. For instance, a meta-analysis of distributed practice showed an inverse relationship between the size of the distributed practice effect and the task's conceptual difficulty (Donovan & Radosevich, 1999). Thus, distributed practice benefits for arithmetic problems may be attenuated, as they are more conceptually demanding than verbatim verbal recall (Rohrer & Taylor, 2006). It would be useful to provide

guidelines regarding the boundaries between when the timing between learning sessions should be extended and when it has been extended too much, especially with regard to arithmetic practice. Therefore, the present study seeks to reconcile the time course for the optimal learning and retention of arithmetic facts.

Time Pressure

Learning arithmetic facts is a gradual process, and it is important not to prioritize speed over accuracy and comprehension. The *What Works Practice Guide for Assisting Students Struggling with Mathematics* (Fuchs et al., 2021) recommends timed practice as a beneficial approach based on evidence from 27 studies. It suggests that timed activities should be embedded in mathematics interventions once students have been introduced to the given arithmetic concept. Educators may use timed practice to improve arithmetic fluency (Knowles, 2010), including traditional timed tests like "mad minute" worksheets. However, we still have limited knowledge about how time pressure affects long-term retention of arithmetic facts.

Although the effects of time pressure on mathematical cognition are not fully understood, some students may feel stressed when taking timed arithmetic tests (Engle, 2002). It is uncertain if this induced stress hinders or improves learning and retention. Additionally, it is worth considering how students perceive time pressure during arithmetic practice. If it is perceived as a challenge or goal to achieve, time pressure can act as a typical challenge stressor and provide an opportunity for growth, which may enhance performance (Widmer et al., 2012).

The question of whether time pressure induces stress is complex. Daily timed practice drills result in higher posttest scores and more improvements between pretest and posttest scores compared to once-a-week timed practice drills or no practice at all (Knowles, 2010). Although this suggests that daily practice may be more effective than weekly practice, the study did not control for the total number of practices or compare timed and untimed practice, highlighting the need for further investigation.

One study compared timed versus untimed practice as part of a larger tutoring intervention for children identified as at-risk for math learning difficulties. The results showed that while timed practice improved the learning of arithmetic and two-digit calculations, it did not enhance number knowledge and word problems (Fuchs et al., 2013), indicating that the effects of time pressure may not be universal for all math tasks, especially those requiring higher levels of mathematical reasoning. However, the study also suggested that timed practice may be beneficial for basic arithmetic facts such as single-digit addition, but these positive effects were shown only on an immediate posttest. A follow-up study found that although the larger tutoring intervention had overall positive effects over a long time period, the differences between timed and untimed practice faded (Bailey et al., 2018), suggesting that time pressure may benefit short-term learning, but it may not persist.

Another concern about time pressure is that it may induce math anxiety (Boaler, 2014), which has the potential to negatively affect performance by depleting working memory, a vital component of mathematical cognition (Ashcraft & Krause, 2007). However, the interaction between emotion and cognition is complex. Optimal cognitive performance occurs with intermediate emotional arousal, and the Yerkes-Dodson Law details how too much or too little arousal negatively affects performance (Yerkes & Dodson, 1908). Similarly, children intrinsically motivated in math show an inverted-U curvilinear relationship between math anxiety and math performance (Wang et al., 2015), suggesting that moderate levels of anxiety can enhance performance in intrinsically motivated students.

Thus, there are two competing predictions. Timed practice may impair the learning of arithmetic facts because students are forced to learn new material in a stressful environment (Boaler, 2014). Alternatively, timed practice may enhance learning by increasing arousal to optimal levels or by compensating for weaker problem-solving strategies (Fuchs et al., 2013). The present study seeks to inform the debate on the effects of time pressure on the learning and retention of arithmetic facts.

The Current Study

The Interaction Between Distributed Practice and Time Pressure

Although distributed practice and time pressure are well-studied factors that are known to affect learning and memory, little research has investigated how these factors interact during math learning. Mathematics achievement depends

on various factors, including domain-specific concepts, skills, and strategies such as number sense and arithmetic decomposition (Busch et al., 2015). Given that multiple factors impact learning in the classroom, it is important to examine their interaction to gain a better understanding of basic learning processes and to provide more realistic and practical guidance for educators.

Our experiment tested competing theories of optimal distribution and the effectiveness of timed versus untimed practice. While numerous studies have focused on identifying the optimal learning conditions for retaining information in the verbal domain, relatively few studies have explored the conditions that promote mathematical cognition. It is possible that the relationship between inter-stimulus and retention intervals may vary depending on the type of memory formed during encoding. Furthermore, few distributed practice studies have investigated time pressure, so analyzing their interaction could offer practical insights into the contexts that support arithmetic learning and retention.

A Priori Hypotheses

We had two competing hypotheses in relation to distributed practice: the exact time distribution hypothesis and the variational distribution hypothesis. With the exact time distribution hypothesis, practice within a week (daily, every other day, weekly) should be superior to the other distribution conditions for recall (every 10 days, every other week) in all post-learning tests. This hypothesis is based on the idea that the 17- and 19-times tables, and arithmetic in general, are not used daily, making them weaker knowledge, and weaker memories fade if not refreshed prior to being lost in the hippocampus. This fading from the hippocampus is predicted to occur at around one week for simple facts presented once (Fisher & Radvansky, 2018), though this time frame may differ for complex materials presented multiple times.

Alternatively, according to the variational distribution hypothesis, daily practice may be more effective than other conditions (every other day, weekly, every 10 days, every other week) for immediate test recall, while every-other-day practice may be superior to other conditions for ten-day delayed test recall. This theory also predicts that the ratio between retention interval and inter-study interval should influence the end-of-semester performance, but the ideal ratio is still unknown. Although longer retention intervals often require longer interstudy intervals, the distribution of time between sessions does not follow a fixed proportion of time between learning and test (Cepeda et al., 2008). Rather, the proportion relating the retention interval and the inter-study interval is variable.

For time pressure, we expected practice with time pressure to be superior to practice without, but we also expected this effect to fade (e.g., its effect should be most noticeable in the immediate test, less noticeable in the ten-day delayed test, and least noticeable in the end-of-semester test). Time pressure would help make memories stronger in the moment, reducing reliance on hippocampal storage. The rationale was that stronger memory may reduce the urgency of including more practice prior to when memories of less frequently used knowledge like arithmetic may fade from the hippocampus at around one week (Squire et al., 2015). Based on math anxiety and math performance's inverted-U curvilinear relationship, if time pressure induces a moderate level of anxiety, it should encourage deeper encoding and automaticity. According to this view, time pressure should enhance arithmetic learning and retention. Based on the effects of timed arithmetic practice seen in Fuchs et al. (2013) and Bailey et al. (2018), we expected mostly short-term effects visible in the immediate test.

Based on our synthesis of the literature, we predicted that the exact time spacing hypothesis would be supported. If so, we further predicted that the effect of condition (within a week vs. after a week) would be reduced with time pressure. Time pressure may create a sense of urgency during learning that may compensate for the longer gap between sessions for participants in the every-10-days and every-other-week conditions. Time pressure may also be less effective for shorter distribution conditions because participants may already have a similar sense of urgency simply by knowing that their next learning session is approaching. Therefore, we reasoned that time pressure may have a larger effect in the longer conditions than in the shorter conditions.

Method

Participants

We powered the study to detect a medium effect size because Davoli and colleagues (2020) found a medium effect of different study conditions in adults learning unfamiliar times tables, and a medium effect is practically meaningful in education. Thus, two hundred and fifty adults were recruited to detect a medium effect size at power = .80 while accounting for an anticipated 20% attrition due to the study's longitudinal nature. The actual attrition was 13.6% for 1-day testers, 11% for 2-day testers, 26.8% for 7-day testers, 24.4% for 10-day testers, 35% for 14-day testers. The final sample for the main analysis included 211 adults (ages 18-40; M_{Age} = 25.2, SD_{Age} = 5.9; 36.5% women and 63.5% men). Descriptive data of education, math anxiety, socioeconomic status, ethnicity, occupation, gender, math interest, age, and current residence are presented in the Appendix. Adults were recruited for the study because, although multiplication is a third-grade standard, second graders' arithmetic proficiency prior to learning multiplication is highly variable (De Smedt et al., 2009). As the first study of this issue, working with adults was also easier than working with children who may or may not have learned the basic concept of multiplication (Imbo et al., 2011), a baseline variability that could complicate analysis.

A total of 39 participants were initially assessed but excluded from the analysis. Participants were excluded for the following reasons: (a) if they did not participate in the three learning sessions and take all four tests (n = 32), (b) if they reported experiencing internet connection issues (n = 1), (c) if they reported using a calculator (n = 1), (d) if they spent less than one second on a problem for multiple problems (n = 4), or (e) if they answered the embedded attention check problems incorrectly (n = 1). The study was conducted through Prolific, an online platform for reliable data collection from a global cohort of participants used in prior psychology studies (e.g., Callan et al., 2017). Participants were screened for at least ten prior submissions and a 95% or above approval rating before they could participate for $10 an hour. There were no demographic constraints. A $2 bonus was awarded for completing all study parts. Participants were given adequate information and filled out an informed consent form. This research was approved by the University of Notre Dame's Institutional Review Board. Data is publicly available (see Supplementary Materials).

Materials

The materials consisted of the 17- and 19-times tables. In the learning phase, the times tables included multipliers from 1 to 12 in a random order, as these are common in children's timed tests. In the testing phase, 3, 6, 7, 8, 9, 12 were multipliers, as these multipliers take longer to mentally calculate, so it is easier to tell when they are not memorized. These arithmetic facts were used to design the pretest, learning phase problem sets, and tests on Qualtrics. The 17-times table had been used as a set of unknown arithmetic facts in the "Hand Position Affects Performance on Multiplication Tasks" study by Davoli and colleagues (2020). To establish that the results were not dependent on the particular arithmetic practiced (e.g., the 17-times table), the 19-times table added variability (Haverty, 1999). There were no expected differences based on the 17- or the 19-times table. To avoid expertise through prior practice, the times tables had high multiplicands: 17 and 19. Because some countries memorize up to the 16-times table, these practiced facts were above the 16-times table to minimize familiarity. Filler facts (random, unpracticed 7-, 9-, 12-times table facts) were also included in the tests. The 7- and 9-times tables were filler facts because they have the same digits in the ones place as the 17- and 19-times tables while the 12-times table similarly requires double digit multiplication. The pretest was identical in format to the post-tests in the testing phase, and six 17-times table facts, six 19-times table facts, and six filler facts were used.

The problem sets at the end of each learning phase were composed of either the 17- or 19-times table, which were randomly assigned to each participant. The assigned times table (e.g., 17) was the multiplicand, so the format of all the learning phase problem sets was a x b = c (e.g., 17 x 4 = __). The problem sets were designed as multiple-choice questions that included the correct answer and seven other randomly ordered answers based on common mistakes children make: Operand-Related errors (using the incorrect multiplier), Miss 1 errors (the answer is off by one from the correct answer), and Adding Instead of Multiplying errors (Buwalda et al., 2016). The adding instead of multiplying error was not included, as it is mostly seen in children who have not mastered multiplication. For 2-digit by 2-digit

multiplication problems, the Forgetting Place Value errors (e.g., 15 x 12 = 30, 30 + 15 = 45) and Rounding Incorrectly errors (e.g., for 15 x 12: 15 x 10 = 150, 150 + [2 x 12] = 174) replaced other common errors. Answers with the same ones digit (Same Ones Digit errors) were also included. This category was also included for incorrect answers to add more variability. A Miss 1 category was included for incorrect answers to prevent searching for consecutive answers, as only using this error for correct answers could signal that the correct answer was a consecutive number. Multiple-choice was used instead of fill-in-the-blank to minimize possible variability due to typing speed or internet connection.

The State-Trait Anxiety Inventory (STAI; Spielberger et al., 1983) assesses how anxious participants feel while completing a task (e.g., state anxiety) and in general (e.g., trait anxiety). To identify whether participants experienced stress, especially when timed, participants filled out a shortened State-Trait Anxiety Inventory (STAI) after each learning phase. This version contained Items 1, 3, 6, 15, 16, and 17 of the State form of the STAI and had a favorable internal consistency (all r's >.90), making it a reliable and valid instrument (Tluczek et al., 2009). The STAI questions asked participants to indicate how they felt right after each learning phase. Using a five-point Likert scale, they indicated their present feelings for "I feel calm. I feel tense. I feel upset. I am relaxed. I feel content. I am worried." Scoring was reversed for items 1, 15, and 16, as recommended in the manual (Spielberger et al., 1983). Higher scores indicated greater state anxiety, a measure of perceived stress.

After the learning phase, the participants started the testing phase. The tests were identical to the pretest, and they included 18 problems: six 17-times table facts, six 19-times table facts, and six filler facts (unpracticed arithmetic: 7-, 9-, 12-times tables). Each test included embedded attention check statements with simpler arithmetic such as "2 x 2 = __." The questions were presented one at a time in black 12-point Arial font and in a random order to each participant during each test. The interstimulus interval depended on when the participant clicked the arrow to move on but had a limit of 90 seconds.

To evaluate understanding of the practiced facts, a "transfer" test was given after the end-of-semester test. This test included three different forms of "transfer" facts: the practiced multiplicand as the multiplier (e.g., when 17 x 7 is the practiced fact, the test problem would be 7 x 17), the non-traditional "answer = operations" (e.g., __ = 17 x 8), and the division version (e.g., 51 / 17). Three 17-times table, three 19-times table, and three filler transfer problems were used. Three filler traditional multiplication problems (e.g., 7 x 4 = __) were added to prevent signaling the variables of interest. The inclusion of "transfer" facts in this study was modeled after the transfer of knowledge in commutative pairs from a study by Davoli and colleagues (2020), in which undergraduates learned multiplication facts (of the format a x b = c) and transferred their learning to a novel format (Exp. 1: b x a = __) or the same format (Exp. 2: a x b = __). Participants used their devices and Prolific accounts and were asked to not use additional materials (e.g., paper and pencil, calculator, times table).

Task

Participants learned and practiced unfamiliar arithmetic facts from a 17- or 19-times table under different distribution conditions with time pressure as a between-subjects manipulation. The primary task was to learn arithmetic facts based on a randomly assigned 17- or 19-times table. The task consisted of five main phases: a pretest, a learning phase (three sessions), an immediate test (the final task in the third learning phase session), a ten-day delayed test, and a delayed test based on a later fixed date (9 weeks after the study start date; Figure 1).

Figure 1

Equal Learning Schedule With Time in Days Depicted Horizontally
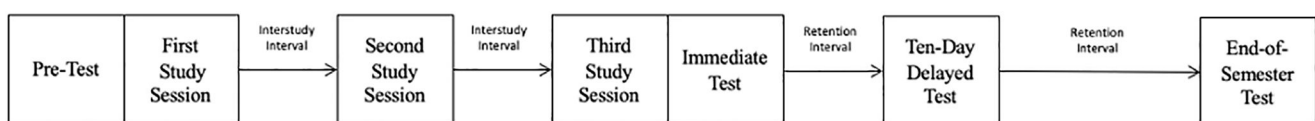
Note. The interstudy interval (ISI) varies with assigned spacing condition: 1, 2, 7, 10, or 14 day intervals. The ten-day delayed test occurs after a fixed retention interval (RI) 10 days after the third study session. The end-of-semester test occurs after a variable RI from the ten-day delayed test (fixed at 9 weeks after the first session).

The pretest, immediate test, and end-of-semester test were the same for participants in all conditions. They mimicked traditional classroom tests with a time limit of 90 seconds per question, but this time limit was not highlighted as a "beat the clock" scenario nor presented to create an acute sense of time pressure regardless of assigned condition.

Procedure

After meeting the criteria, signing up, and providing informed consent, participants began the learning phase on Qualtrics. They read instructions describing the study's phases, how to answer questions, and its purpose. They were reminded of their anonymity. Due to an anticipated wide range in baseline levels of arithmetic proficiency, a pretest was administered at the start of the first learning phase session to gauge mathematical abilities. The pretest was also a baseline measure to ensure that question order did not influence answers and to assess participant learning over time.

Each participant was given the following task during each of the three learning phase sessions: solve a times table, study the times table, and take an assessment. Participants were first asked to solve a 17- or 19-times table with multipliers presented in a random order and to study their assigned times table with consecutive multipliers for as long as they would like, although this was roughly equal across conditions with an average of 6 minutes. When ready to move on, participants were told to try their best to solve each problem as quickly as they could while still maintaining high accuracy. They then recalled the studied arithmetic by filling out two consecutive 20 multiple-choice problem sets. In the timed condition of the learning phase, participants were given the goal of trying to solve at least 15 problems correctly in 30 seconds. After 30 seconds, regardless of how much was completed, each problem was displayed along with a red 0/1 (incorrect) or a green 1/1 (correct) score. Participants could review their answers for as long as they liked before moving on. The next screen showed how many problems they answered correctly and whether they met the goal of correctly answering at least 15 problems in 30 seconds. Any problems answered incorrectly or left blank reappeared to be redone. Participants then had a chance to beat their initial scores with another set of 20 problems with the same goal. Meanwhile, in the untimed condition, people were also given two sets of 20 problems but were not made aware of the time spent. Like the timed condition, answers were scored and revealed. They also reviewed answers, were notified of their accuracy, and re-solved incorrect or unanswered problems.

After completing the first learning phase session tasks, participants completed the State form of the modified State-Trait Anxiety Inventory. As an additional reliability measure, participants then answered a question about whether there was anything they had to share about their participation (e.g., opening a times table in a separate tab, connection issues). They were reminded that their answers would not impact their monetary compensation and completed a survey that collected their demographic information. This self-reported measure was included at the end of data collection to minimize biased responses during the study.

After all three learning phase sessions, participants completed a filler task (the shortened State-Trait Anxiety Inventory) followed by the immediate test. The end-of-semester test was administered 9 weeks after the start of the study. To serve as a more practical measure for how educators could incorporate the study's results into teaching, it was designed as a fixed date rather than a variable ratio with the last learning session. Nine weeks were chosen because K-12 schools in the U.S. commonly use a nine-week grading period system; this is thought to provide more time to learn complex concepts and to receive feedback and intervention prior to grade deadlines despite limited empirical evidence supporting this idea. With more schools shifting to a nine-week grading period, it may be helpful to assess the learning and retention of arithmetic using this grading period system.

Design

The design was a 5 (spacing condition: 1, 2, 7, 10, or 14 day intervals) x 2 (time pressure: timed or untimed) x 2 (fact type: practiced or unpracticed) x 4 (test phase: pretest, immediate, ten-day delayed, end-of-semester) mixed design with repeated measures on fact type and test phase. The distribution of study times was modified based on studies by Fisher and Radvansky (2018) and Cepeda et al. (2008). The spacing effects on the end-of-semester test were extrapolated from Cepeda et al. (2008) data of test performance as a function of the inter-study interval. Participants were randomly assigned to one of the ten between-subjects conditions in the 5 (spacing) x 2 (time pressure) factorial design for the learning phase. Bayesian analyses were performed to directly test the null hypothesis while comparing multiple

competing hypotheses. A Bayes factor of less than 1 indicates evidence for the null hypothesis while greater than 1 indicates evidence for the alternative hypothesis. The larger the Bayes factor's absolute value, the stronger the evidence.

Results

Response Time

Response times (RTs) ranged from 2 to 90 s with a mean RT per problem of 12 s (SE = .2 s) across all conditions because the accuracy of the Qualtrics timing was to the second. The response time data at each test was right-skewed, $W(1806)$ = .79, $p < .001$, with skewness of 2.65 (SE = .06) and kurtosis of 12.96 (SE = .12). This is an expected and typical finding with response time data. Given that ANOVA is robust against violations of normality, we proceeded with the analysis. We did a 5 (spacing condition: 1, 2, 7, 10, or 14 day intervals) x 2 (time pressure: timed or untimed) x 2 (fact type: practiced or unpracticed) x 4 (test phase: pretest, immediate, ten-day delayed, end-of-semester) mixed analysis of variance (ANOVA) with repeated measures on fact type and test phase.

Mauchly's test indicated that the assumption of sphericity had been violated for test phase, $\chi^2(5)$ = 51.70, $p < .001$, therefore, degrees of freedom were corrected using Huynh-Feldt estimates of sphericity (ε = .920) for the within-subjects tests involving test phase. Two main effects were found. First, there was a main effect of test phase, $F(2.76, 532.90)$ = 40.93, $p < .001$, η_p^2 = .18. Test phase (pretest, immediate test, ten-day delayed test, end-of-semester test) had an effect on the response time during recall (Figure 2). Planned orthogonal Helmert contrasts revealed that the performance on the pretest was slower than the three posttests, $F(1, 193)$ = 78.12, $p < .001$. The immediate test was faster than the ten-day and end-of-semester tests, $F(1, 193)$ = 35.93, $p < .001$, but the ten-day and end-of-semester tests did not differ, $F(1, 193)$ = .08, p = .78. Second, there was a main effect of fact type (practiced or unpracticed), $F(1, 193)$ = 62.82, $p < .001$, η_p^2 = .25. Responses were faster on the practiced (M = 11, SE = .3) than on the unpracticed facts (M = 13, SE = .4).

Figure 2

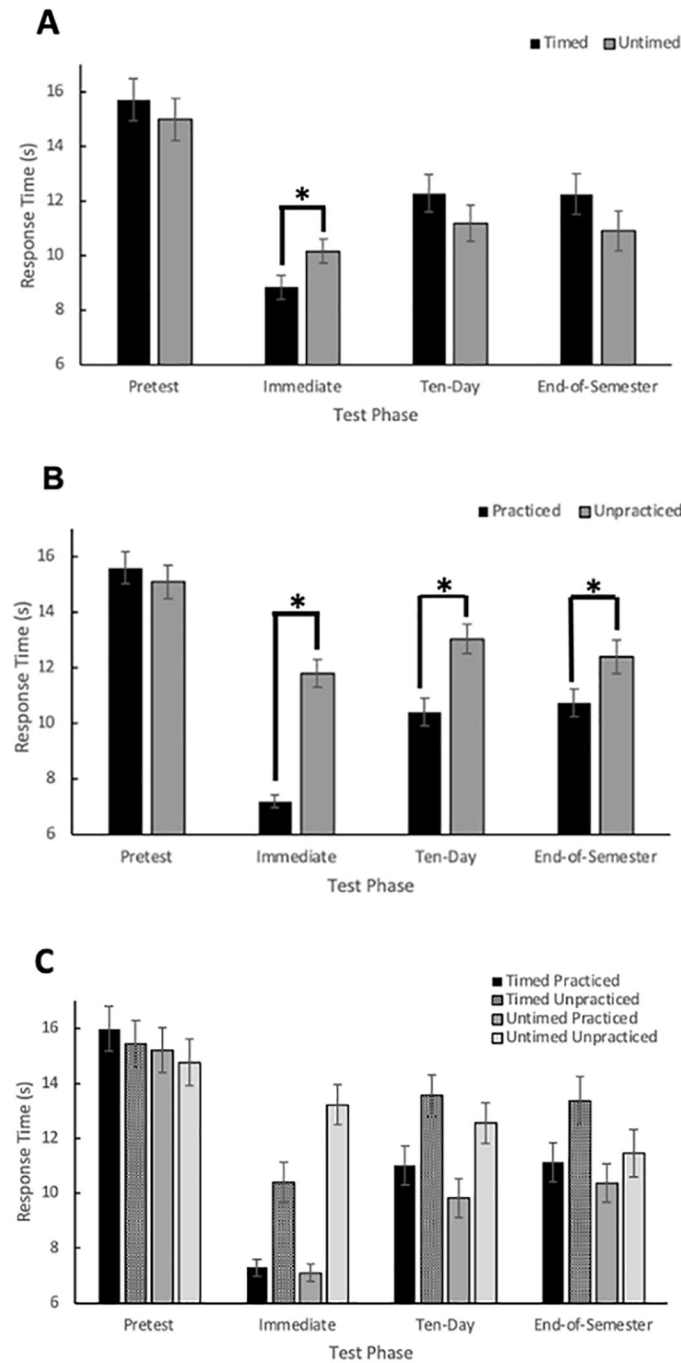*Effect of Test Phase on the Response Time for Arithmetic Fact Recall*
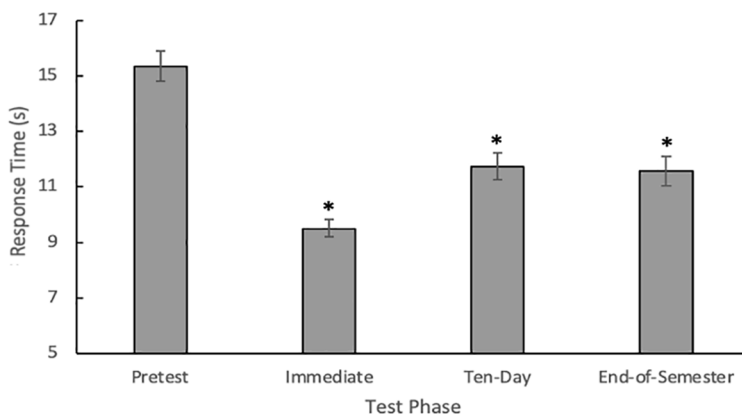
Note. Response time on a pretest and three posttests was assessed among participants (n = 211) in all spacing and timing conditions. Marginal means ± standard error of the mean.
*$p < .05$ as compared to pretest.

Beyond the main effects, there were significant test phase x fact type, $F(3, 579)$ = 30.18, $p < .001$, η_p^2 = .14 (Figure 3B) and test phase by fact type x time pressure interactions, $F(3, 579)$ = 5.16, p = .002, η_p^2 = .03 (Figure 3C). The untimed learning condition gradually improved in response time on unpracticed facts across the test phases (M = 14.8, SE = .8; M = 13.2, SE = .7; M = 12.6, SE = .7; M = 11.4, SE = .9) while the other conditions showed the more typical forgetting after initial improvement on the immediate test. The test phase by time pressure interaction, $F(2.76, 532.90)$ = 2.56, p = .059, η_p^2 = .01, was not significant (Figure 3A).

Figure 3

Effects of Time Pressure (A) and Fact Type (B) Across Test Phase and the Interaction Between Test Phase, Fact Type, and Time Pressure (C) on the Response Time of Arithmetic Fact Recall

Note. Response time on a pretest and three posttests was assessed among participants (n = 211) in all spacing conditions and compared between the timed (time pressure, n = 104) and untimed (no time pressure, n = 107) conditions (A) as well as for practiced facts (e.g., 17-times table facts for participants assigned to the 17-times table) versus unpracticed facts (e.g., 19-times table facts for participants assigned to the 17-times table) (B). Response time on a pretest and three posttests was then compared between timed and untimed condition participants and their performance on practiced versus unpracticed facts (C). Marginal means ± standard error of the mean.
*p < .05.

However, Post-Hoc analysis indicated that time pressure improved performance on the immediate test (M = 8.84, SE = .45; M = 10.16, SE = .44). The other three-way interactions and the four-way interaction were all not significant (all p > .05).

A Priori Hypotheses

We assessed response time in relation to our two hypotheses. Contrary to expectations, we found no main effect of distribution, $F_{(4, 193)}$ = .16, p = .96, η_p^2 = .003, BF_{10} = .03. Similarly, there was no main effect of time pressure, $F_{(1, 193)}$ = .46, p = .50, η_p^2 = .002, BF_{10} = .20. However, there was a spacing condition x time pressure interaction, $F_{(4, 193)}$ = 3.44, p = .01, η_p^2 = .07 (Figure 4). This two-way interaction was qualified by the three-way interaction, $F_{(3, 579)}$ = 5.16, p = .002, η_p^2 = .03 (Figure 5). We did not test our final *a priori* hypothesis of the spacing effect's reduction in the time pressure condition, as it depended on a main effect of distribution on response time.

Figure 4

Effect of the Interaction Between Time Pressure and Spacing Condition on the Response Time of Arithmetic Fact Recall
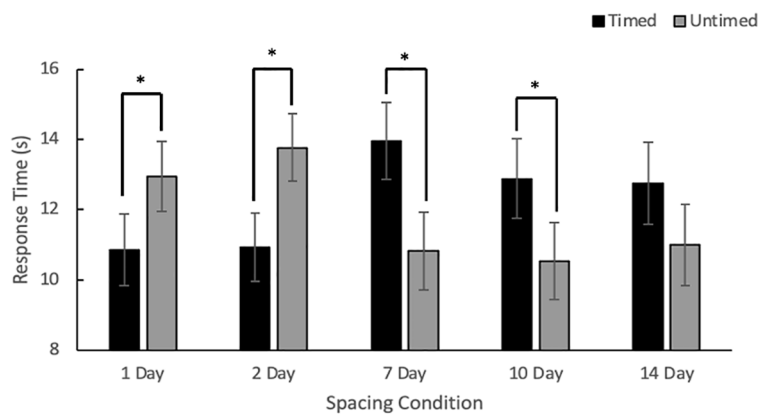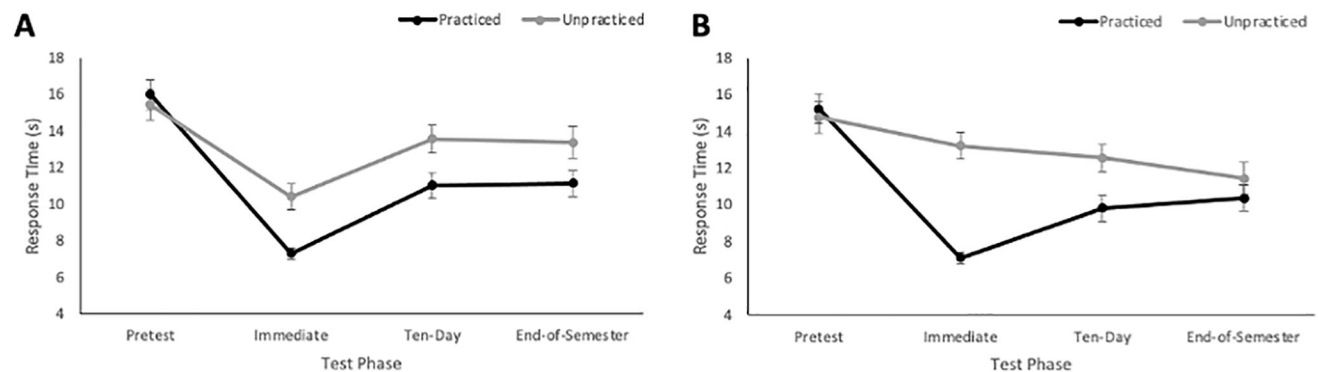

Note. Response time on a pretest and three posttests was assessed among participants (n = 211) in all spacing conditions and compared between the timed (time pressure, n = 104) and untimed (no time pressure, n = 107) conditions. Marginal means ± standard error of the mean.
*p < .05.

Figure 5

Effect of the Three-Way Interaction Between Test Phase, Fact Type, and Time Pressure on the Response Time of Arithmetic Fact Recall

Note. Response time on a pretest and three posttests was compared between the timed condition participants (A) and the untimed (B) and their performance on practiced facts (e.g., 17-times table facts for participants assigned to the 17-times table) versus unpracticed facts (e.g., 19-times table facts for participants assigned to the 17-times table). The time between the ten-day and end-of-semester posttests varies according to spacing condition (producing either an expanding or contracting schedule), which would warrant a bar graph. However, the relationship across test phase (where time is a continuous variable) was best demonstrated with a line graph. The untimed condition participants' performance on unpracticed facts followed a different forgetting trajectory from the other fact type by timing interaction conditions. Marginal means ± standard error of the mean.

Robustness Checks

To further understand the test phase x fact type, three-way, and the distribution condition x time pressure interactions, we re-analyzed the data without the pretest to see if the interactions still held, and they did in all posttests. We also checked if the interactions held when controlling for average accuracy as a covariate. This showed that the test phase x fact type interaction no longer held. However, the test phase x fact type x time pressure interaction did, as did the distribution condition x time pressure interaction.

Exploratory Analysis

Because the two post-practice interactions were not predicted *a priori*, we did an exploratory analysis. The first major finding was that timed practice decreased response time for participants in the 1- and 2-day conditions while increasing response time for participants in the 7-, 10-, and 14-day conditions (Figure 6). This interaction's effect size was medium, $F(1, 199) = 13.60$, $p < .001$, $\eta_p^2 = .06$.

Figure 6

*Effects of Time Pressure and Within a Week Versus Greater Than a Week Spacing Conditions on Response Time*
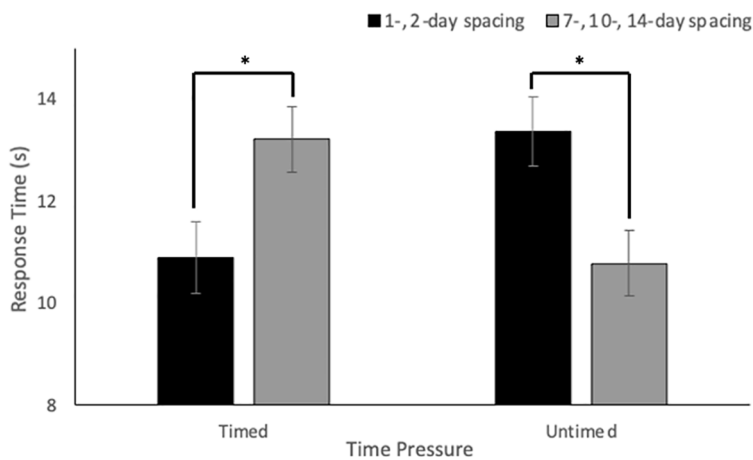
Note. Response time was assessed according to spacing condition (within a week, $n = 89$; or, greater than a week, $n = 122$) and timing condition (timed practice, $n = 104$; untimed practice, $n = 107$) for all tests excluding the pretest. Estimated marginal means ± standard error of the mean.
*$p < .05$.

The second major finding was that the difference between the practiced and unpracticed facts seemed to follow the same forgetting trajectory across the test phase among timed condition participants (Figure 5; Figure 3C). This trend was not found for untimed condition participants. Rather, these participants gradually improved on the unpracticed facts with repeated practice across the posttests.

Accuracy

We anticipated near-ceiling performance in accuracy. Accuracy ranged from .00 to 1.00 with a mean of .84 ($SE = .01$) across all conditions. The accuracy at each test was left-skewed, $W(1780) = .74$, $p < .001$, with skewness of -1.610 ($SE = .06$) and kurtosis of 2.191 ($SE = .12$). Because ANOVA is robust against violations of normality, we proceeded with the analysis. Participants exhibited more variability across test phase than expected on the arithmetic recall task (Figure 7). Seeing condition effects in accuracy was a novel finding, so we moved to testing our hypotheses using accuracy data. We did a 5 (distribution condition: 1, 2, 7, 10, or 14 day intervals) x 2 (time pressure: timed or untimed) x 2 (fact type: practiced or unpracticed) x 4 (test phase: pretest, immediate, ten-day delayed, end-of-semester) mixed analysis of variance (ANOVA) with repeated measures on fact type and test phase.

Figure 7

Effect of Test Phase on the Accuracy of Arithmetic Fact Recall
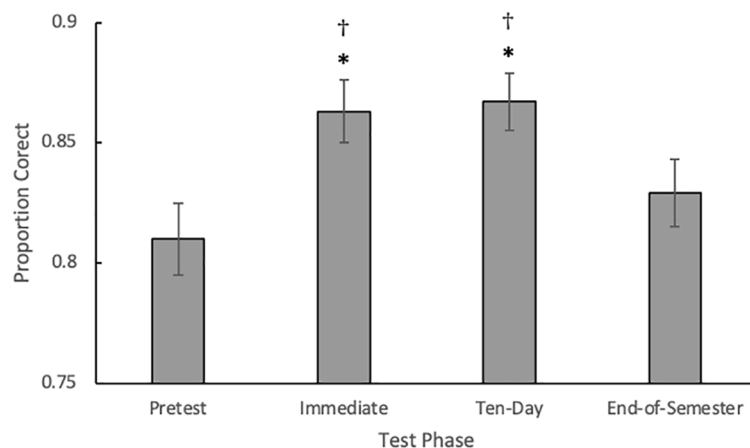
Note. Accuracy on a pretest and three posttests was assessed among participants (n = 211) in all spacing and timing conditions. Marginal means ± standard error of the mean.
$^*p$ < .05 as compared to pretest, $^†p$ < .05 as compared to end-of-semester test.

Mauchly's test indicated that the assumption of sphericity had been violated for test phase, $\chi^2(5)$ = 70.58, p < .001. Degrees of freedom were corrected using Huynh-Feldt estimates of sphericity (ε = .902) for the within-subjects tests involving test phase. Two main effects were found. First, there was a main effect of test phase, $F(2.71, 503.31)$ = 8.77, p < .001, η_p^2 = .045 (Figure 7). To mirror response time data analysis, we next moved to the orthogonal Helmert contrasts. The Helmert contrasts revealed that accuracy at pretest was worse than the average accuracy on the three posttests, $F(1, 186)$ = 13.20, p < .001. The immediate test was not more accurate than the ten-day delayed and end-of-semester tests, $F(1, 186)$ = 2.62, p < .11, but the ten-day delayed test was more accurate than the end-of-semester test, $F(1, 186)$ = 7.59, p = .006. Second, the main effect of fact type (practiced or unpracticed) was significant, $F(1, 186)$ = 32.94, p < .001, η_p^2 = .15. Accuracy was higher on the practiced facts (M = .87, SE = .01) than on the unpracticed facts (M = .82, SE = .01).

These main effects were qualified by significant test phase by time pressure, $F(2.71, 503.31)$ = 3.67, p = .015, η_p^2 = .019, test phase by fact type, $F(3, 558)$ = 11.79, p < .001, η_p^2 = .060, interactions. The timed condition was less accurate than the untimed condition on the immediate test (Figure 8A). People were also more accurate on the practiced facts than on the unpracticed facts for the immediate and ten-day tests (Figure 8B). This was not the case for the other tests. The other three-way interactions and the four-way interaction were not significant (all p > .05).

A Priori Hypotheses

We assessed accuracy in relation to our two hypotheses. We did not find evidence for a distribution effect on accuracy, $F(4, 186)$ = 2.0, p = .10, η_p^2 = .04, BF_{10} = .14. We also did not find evidence that timed practice led to higher accuracy than untimed practice. The means were in the opposite direction of our prediction, although the difference was not significant, $F(1, 186)$ = 1.98, p = .16, η_p^2 = .01, BF_{10} = 3.23. Again, we did not test the hypothesis that the distribution effect would be reduced in the time pressure condition, as it depended on finding a main effect of distribution.

Transfer Test

We conducted a 5 (spacing condition: 1, 2, 7, 10, or 14 day intervals) x 2 (time pressure: timed or untimed) x 2 (fact type: practiced or unpracticed) mixed analysis of variance (ANOVA) with repeated measures on fact type using the transfer test accuracy and response time data. Accuracy results revealed a main effect of fact type, $F(1, 207)$ = 5.52, p = .02, BF_{10} = 2.11, with it being higher for practiced (M = .86, SE = .02) than unpracticed facts (M = .81, SE = .02). A main effect of fact type on response time was indeterminate, $F(1, 207)$ = 1.89, p = .17, BF_{10} = .48. A main effect of timing condition was

found on accuracy, $F(1, 207) = 4.12$, $p = .04$, with more accurate responses in the timed than untimed condition ($M = .87$, $SE = .02$; $M = .80$, $SE = .02$, respectively). This was not seen for response time. There were no main effects of distribution on accuracy or response time.

Figure 8

Effects of Time Pressure (A) and Fact Type (B) Across Test Phase
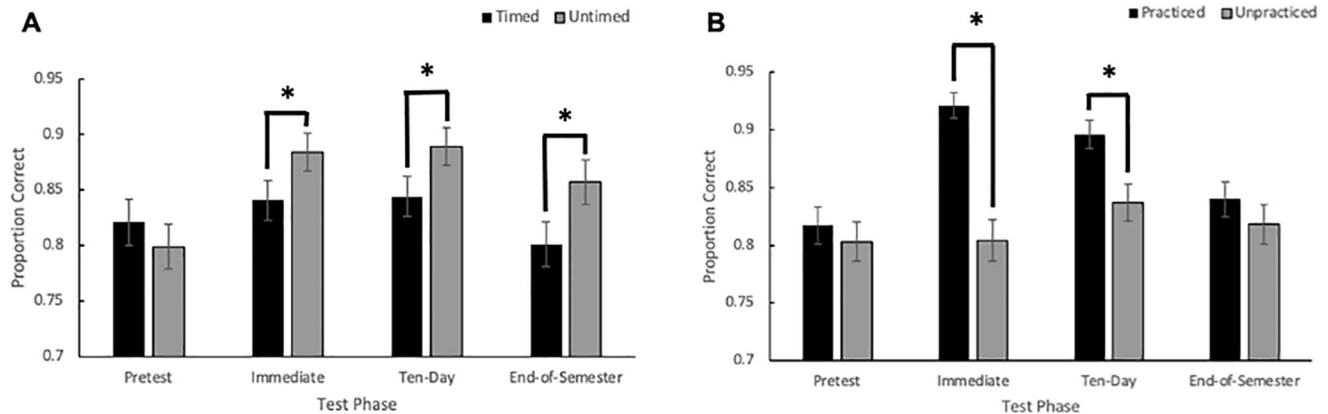
Note. Accuracy on a pretest and three posttests was assessed among participants ($n = 211$) in all spacing conditions and compared between the timed (time pressure, $n = 104$) and untimed (no time pressure, $n = 107$) conditions (A), for practiced facts (e.g., 17-times table facts for participants assigned to the 17-times table) versus unpracticed facts (e.g., 19-times table facts for participants assigned to the 17-times table) (B), and the interaction between time pressure and fact type across the test phases (C). Marginal means ± standard error of the mean.
*$p < .05$.

Stress Measure

To determine the impact of time pressure during learning, the State-Trait Anxiety Inventory (STAI) was given after each learning phase session. Using state anxiety ratings (STAI-state), we did a 2 (time pressure: timed or untimed) x 3 (learning phase session: first, second, or third learning phase) mixed analysis of variance (ANOVA) with repeated measures on learning phase session. There was a main effect of timing condition on reported state anxiety, $F(1, 10) = 14.81$, $p = .003$, $\eta_p^2 = .60$, a main effect of learning phase session on reported state anxiety, $F(2, 20) = 58.65$, $p < .001$, $\eta_p^2 = .85$, and an interaction between timing condition and learning phase session, $F(2, 20) = 28.27$, $p < .001$, $\eta_p^2 = .74$. Overall, participants in each time pressure condition reported experiencing less stress with each subsequent session. Participants in the timed condition reported experiencing more stress than those in the untimed condition following all sessions. They reported experiencing similar stress during the first and second learning sessions, and it was not until the third session that their stress decreased, while the participants in the untimed condition reported experiencing less stress during the second learning session (Table 1).

Common Arithmetic Errors

We did an exploratory analysis for the most common arithmetic error by recording how often a type of error appeared among incorrect answers for different fact types (17- or 19-times table) across test phase (Pretest, Immediate, Ten-Day, End-of-Semester) and the overall frequency of each error (Table 2). Recall that the tests were multiple choice, so errors were limited to the answers we had included as foils based on previous research. We had an exploratory hypothesis (see pre-registration in the Supplementary Materials) that testing pressure could lead participants who relied on calculation rather than retrieval to hastily select an answer once the correct ones digit was known, resulting in the "Same Ones Digit" error. This error was indeed the most common (37%). Overall, 162 of the 211 participants (77%) made the Same Ones Digit error at least once during testing. The next most frequent error was Rounding Incorrectly (23.7%), which is only found in 2-digit by 2-digit multiplication. Place Value errors were non-existent after the pretest.

Table 1

Means of Shortened STAI Scores for Timed Versus Untimed Participants Following Each Learning Phase

	Learning Phase 1		Learning Phase 2		Learning Phase 3	
Item	Timed	Untimed	Timed	Untimed	Timed	Untimed
1 (calm)	3.18	2.88	3.17	2.32	2.82	2.15
3 (tense)	3.61	3.19	3.55	2.67	3.32	2.50
6 (upset)	2.80	2.39	3.07	1.87	2.80	1.91
15 (relaxed)	3.52	3.18	3.42	2.71	3.24	2.45
16 (content)	3.19	2.82	3.18	2.33	2.97	2.40
17 (worried)	2.77	2.45	2.83	1.96	2.62	2.11
M	3.18	2.82	3.20	2.31	2.96	2.25
SEM	0.14	0.14	0.10	0.14	0.11	0.09

Note. Overall mean and SEM reported. Scoring was reversed for items 1, 15, and 16, as recommended in the STAI scoring manual. Higher scores indicate greater state anxiety, which is how anxious participants felt right after completing each learning phase session.

Table 2

Percentage of Participants Who Made Different Types of Arithmetic Errors Across Test Phase

		Percentage of Arithmetic Errors per Error Type (%)						
Test	Fact Type	Operand	Miss 1	Place Value	Rounding	Same Ones	Miss 1 (IC)	Same Ones (IC)
Pre	17	3.7	11.6	3.0	31.5	28.0	7.5	14.7
	19	4.6	13.0	1.9	14.9	52.6	8.5	4.5
Immediate	17	13.7	9.9	0.0	17.0	31.3	6.3	21.8
	19	14.1	16.9	0.0	3.4	48.0	10.2	7.4
Ten-Day	17	5.3	7.3	0.0	34.7	23.1	6.9	22.7
	19	6.8	17.9	0.0	9.4	51.0	10.6	4.4
End	17	5.4	5.2	0.0	38.6	27.3	8.0	15.5
	19	8.0	12.0	0.0	20.8	42.3	10.8	6.2
Overall Percentage (%)		7.0	10.7	0.8	23.7	37.0	8.5	12.3

Note. The arithmetic errors reported include Operand-Related errors (using the incorrect multiplier), Miss 1 errors (the answer is off by minus one or plus one), Forgetting Place Value errors (only for 2-digit by 2-digit; e.g., 15 x 12 = 30, 30 + 15 = 45), Rounding Incorrectly errors (only for 2-digit by 2-digit; e.g., for 15 x 12: 15 x 10 = 150, 150 + [2 x 12] = 174), and Same Ones Digit errors. To prevent searching for consecutive answers and to add more variability to the answer choices, a Miss 1 category and a Same Ones Digit category were added for incorrect answers (IC). The overall percentage for how often each error type was selected is also displayed.

Discussion

The current study examined how distributed practice and time pressure affect the learning and retention of unfamiliar arithmetic facts. Our goal was to contribute to the ongoing search for optimal ways to structure practice. The study was designed to extend prior knowledge on distributed practice by integrating time pressure and considering effects on mathematical cognition. The main effects of test phase and fact type increased our confidence in the major findings. Participants took more time and were less accurate on the pretest compared to the immediate posttest. However, they took more time and were less accurate as the retention interval increased, suggesting that learning was followed by forgetting (Figures 2 and 7). Participants performed better on practiced than unpracticed facts. Despite not finding evidence in support of our *a priori* hypotheses, results revealed the following: (a) time pressure elevated perceived stress

and induced state anxiety, (b) distributed practice and time pressure interacted: time pressure during learning led to faster test performance if the inter-stimulus study interval was less than one week but slower test performance for longer inter-stimulus study intervals, (c) the presence or absence of time pressure contributed to a different forgetting trajectory of practiced and unpracticed facts, and (d) learning transferred to conceptually related facts.

Time Pressure as a Stressor

Participants who were timed during learning scored higher on the state anxiety portion of the State-Trait Anxiety Inventory (Table 1). The state anxiety component of the inventory assesses how anxious participants feel during learning, and our results suggest that time pressure increased participants' perceived anxiety levels. Timed participants reported feeling subjectively more stressed than those in the untimed condition. Although future studies may benefit from using math-specific anxiety measures, prior research on time pressure supports our findings. Time pressure has been shown to heighten the stress associated with decision-making tasks (Byrne et al., 2015; Young et al., 2012). Despite the increased stress reported by timed participants, there were no main effects of time pressure during learning on test performance. Nevertheless, the timed problem sets were perceived as stressful and affected performance differently across different spacing conditions.

The Interaction Between Spacing and Time Pressure

The second major finding was an interaction between time pressure and time between study sessions. Participants in the timed learning phase who studied material daily or every other day responded faster during the testing phase as compared to participants in the untimed learning phase. However, this pattern reversed when the inter-stimulus interval was around one week. Participants in the untimed learning phase who studied material weekly, every 10 days, or every other week responded faster during the testing phase than those in the timed learning phase (Figures 4 and 6). Therefore, time pressure during learning was found to enhance response times during testing only when paired with closely spaced together practice sessions. This finding is consistent with prior studies in which daily timed practice enhanced performance as compared to farther-spaced timed practice or no practice (Duhon et al., 2022; Knowles, 2010), although these prior studies did not compare timed practice with untimed practice.

The hippocampus plays a critical role in memory (Scoville & Milner, 1957). Hippocampal memory loss is thought to occur approximately one week after first learning new material (Fisher & Radvansky, 2018). A hippocampal explanation would initially predict a main effect of spacing, as this has been empirically demonstrated in the memory literature (Hintzman et al., 1973; Kim et al., 2019; Shaughnessy et al., 1972). However, this robust and reliable finding has not been applied to arithmetic learning. Although the current study did not find a main effect of practice distribution, we observed an interaction between time pressure and distribution. One possible explanation for this is that time pressure did not compensate for inter-stimulus intervals longer than one week when learning arithmetic. Compared to language, arithmetic is less frequently used knowledge, largely because it is often disliked and avoided outside of educational settings (Cannon & Ginsburg, 2008). Although time pressure may create a beneficial sense of urgency (Squire et al., 2015), our study suggests that timed practice during learning led to faster response times during testing only when memories were recently consolidated. Therefore, the effects of time pressure may depend on the temporal dimensions of hippocampal-neocortical dialogue, during which hippocampal-dependent memories stabilize in the cortex (Sweegers et al., 2014), allowing for the transition of problem-solving strategies to more automatic retrieval (Rickard, 1997), and eventually leading to skill learning (Rickard et al., 2008). Although the precise temporal dimensions of this process are unknown (Squire et al., 2015), our results suggest that it likely occurs during an inter-stimulus interval between two and seven days.

Another possibility is that the longer inter-stimulus intervals functioned as stressors because a lack of recent review may be perceived as stressful. Time pressure is another stressor, as corroborated by the shortened State-Trait Anxiety Inventory (Table 1). By adding timed practice, the stressors may accumulate to surpass the amount of beneficial stress, as depicted by the inverted-U-shaped relationship between spacing and free recall (Verkoeijen et al., 2008). This may occur at an inter-stimulus interval between two and seven days, contributing to the observed time pressure and spacing trend reversal.

The Effects of Time Pressure on Forgetting

Our third major finding indicated that the forgetting trajectory differed based on whether practice occurred under time pressure, as evidenced by the time pressure, fact type, and test phase interactions. In the timed condition, both practiced and unpracticed facts followed a typical learning-then-forgetting-over-time pattern across the multiple test phases. Similarly, practiced facts in the untimed condition followed this pattern. However, unpracticed facts in the untimed condition seemed to allow participants to gradually improve on unpracticed facts across the tests (Figures 5 and 3C). The forgetting that occurred for the unpracticed facts in the timed condition suggests that the results are not merely a general effect of additional practice during the testing phase.

Transfer Test

Finally, we found that participants who learned a specific set of arithmetic facts were able to apply their knowledge to solve similar problems, indicating transfer of learning. Accuracy was higher for transfer facts that used the same times table as the practiced facts, suggesting that participants constructed conceptual knowledge. These results build upon a previous study where undergraduates learned multiplication facts (*a* x *b* = *c*) and then were asked to either transfer their learning to a novel format (*b* x *a* = __) or the same format (*a* x *b* = __) (Davoli et al., 2020). Our findings support the conclusion that learning occurred, as knowledge that is not transferred to new contexts or problems is considered inert (Bransford et al., 2001). While retention requires recall, transfer requires understanding and application (Bransford et al., 2001; Haskell, 2001). Additionally, the lack of time pressure in the untimed condition may have contributed to an enhanced conceptual understanding, possibly through a relational strategy like decomposition, which is associated with accurate equation encoding (Chesney et al., 2014). Alternatively, participants may have relied on operational patterns, as prior practice did not affect response times. Operational patterns may be helpful for traditional arithmetic, but are unhelpful when encoding, interpreting, and solving non-traditional problems (McNeil et al., 2015). Overall, although participants showed enhanced learning of the practiced facts, as seen in the accuracy data, the transfer of learning to a novel format takes time.

Implications

The present results demonstrate that time pressure during learning and previously practiced arithmetic facts affect test performance. While forgetting occurred for the other conditions, participants gradually improved on unpracticed facts with untimed practice, albeit with slower response time on the immediate posttest and no difference on the end-of-semester posttest compared to other groups (Figure 5). Although it is unclear whether performance would continue to improve with a longer retention interval, reviewing previously unlearned but related information in an untimed, stress-reduced environment may allow students to learn from tests. One way to implement such an environment is through something like math story time, which can increase children's mathematics achievement (Berkowitz et al., 2015). Given that time pressure only helps if it is paired with less spaced apart review (e.g., 1- or 2-day spacing), educators should consider the length of time between study sessions when using timed practice. These findings provide insights for optimizing arithmetic practice and using limited instructional time effectively.

The current study also contributes to the theoretical understanding of how the time between study sessions and the presence of time pressure influence learning and memory. It advances knowledge by examining memory for mathematical information, an area that is studied less than verbal information in memory research. The study did not replicate the typical distributed practice effect seen in the memory literature. This lack of replicability suggests that the ratio between the inter-stimulus interval and retention interval may vary based on the type of material studied, such as arithmetic, narrative texts, or word lists. Furthermore, the study provides insights on the relationship between practiced facts and the transfer to conceptually related facts across different interstudy intervals and time pressure conditions. Participants responded more accurately but not faster on practice facts, and untimed participants were more accurate, which deepens our understanding of how these factors affect learning.

Limitations

Despite the current study's theoretical and practical contributions, it is not without limitations. First, the 1-, 2-, and 7-day spacing conditions were in the "within a week" group, but each learning session was open on Prolific for 1-2-, 2-3-, and 7-8-days. Given these classifications, the 7-day group could have been part of the "greater than a week" category. That said, the 1-week time is a rough estimate. Second, the end-of-semester test, although practically relevant, may be theoretically uninterpretable because of the differences in the retention interval between the groups. Our study used an expanding, contracting, and equal learning schedule where the inter-stimulus interval respectively increased, decreased, or remained constant. An inconsistent pattern of learning episodes contributed to a variable retention interval between the ten-day delayed and end-of-semester tests, possibly acting as a confound for end-of-semester test interpretability, as the optimal learning schedule may depend on the retention interval (Küpper-Tetzel et al., 2014). Despite this, the end-of-semester test is ecologically relevant because the testing situation closely resembles a real class: assessing what students know, teaching the content and standards, testing them shortly after learning to assess what they retained, and administering an end-of-semester assessment. Lastly, practicing just one times table (17 or 19) makes our task somewhat different from a more natural learning situation where there would be higher potential for cross-times table interference.

Future Directions

The current study sheds light on how timed practice and distributed practice affect learning and retention of arithmetic facts, which may differ from the learning and retention of other information (e.g., narrative texts, word lists). More research is needed to understand the potential differences and interactions between time pressure and inter-stimulus intervals of less than one week versus greater than one week. One avenue for future work is to conduct a study with multiple inter-stimulus intervals between two and seven days to identify the exact time when practice with time pressure is no longer preferable. Understanding how spacing and time pressure impact participants' retrieval comfort level is also worth exploring, as individuals differ greatly in their retrieval comfort level. Retrieval comfort refers to how comfortable individuals are with relying on retrieval to solve arithmetic problems (Hecht, 2006; Siegler, 1988). For example, time pressure may harm learning for "not-so-good retrievers," who are slower at executing fact retrieval processes, but may have a smaller effect on "good retrievers" who consistently opt for retrieval even when strategy choices are allowed. The impact on "perfectionists," who are reluctant to rely solely on retrieval despite retrieving correct answers, is difficult to predict. Time pressure could serve as a form of exposure to help them overcome their discomfort by forcing them to practice retrieval, or it could lead to increased anxiety and reduce their retrieval comfort level further in the long run.

It is also important for future studies to address the end-of-semester test's theoretical interpretability by excluding the immediate and ten-day delayed tests, as they may act as learning sessions. To deepen our understanding of the interaction between spacing and timed practice, future research should systematically examine mechanisms that contribute to the gradual improvement on unpracticed facts that occurs in the untimed learning condition. If we can ascertain why the relationship between time pressure and spacing reverses between two and seven days, and why a lack of time pressure results in gradual improvement on unpracticed facts, then we will have a fuller picture of how distributed practice and time pressure affect math performance. Ultimately, this line of work has the potential to assist mathematics educators in making informed and nuanced decisions about how best to allocate their valuable instructional time to optimize student learning of arithmetic facts.

Supplementary Materials

The Supplementary Materials contain the following items (for access see Index of Supplementary Materials below):

- the pre-registration protocol for the study
- the research data supporting the findings

Index of Supplementary Materials

Jazbutis, O. R., & McNeil, N. M. (2020). *Memory and arithmetic facts: Examining the effects of spacing and time pressure on the learning and retention of arithmetic facts* [Pre-registration protocol]. OSF Registries. https://doi.org/10.17605/OSF.IO/2BC7P

Jazbutis, O. R., & McNeil, N. M. (2023). *Supplementary materials to "Distributed practice and time pressure interact to affect learning and retention of arithmetic facts"* [Research data]. OSF. https://osf.io/vnwz6

References

Anderson, J. R. (2002). Spanning seven orders of magnitude: A challenge for cognitive modeling. *Cognitive Science, 26*(1), 85–112. https://doi.org/10.1207/s15516709cog2601_3

Arroyo, I., Royer, J. M., & Woolf, B. P. (2011). Using an intelligent tutor and math fluency training to improve math performance. *International Journal of Artificial Intelligence in Education, 21*(1-2), 135–152. https://doi.org/10.3233/JAI-2011-020

Ashcraft, M. H., & Krause, J. A. (2007). Working memory, math performance, and math anxiety. *Psychonomic Bulletin & Review, 14*(2), 243–248. https://doi.org/10.3758/BF03194059

Bailey, D. H., Fuchs, L. S., Gilbert, J. K., Geary, D. C., & Fuchs, D. (2018). Prevention: Necessary but insufficient? A 2-year follow-up of an effective first-grade mathematics intervention. *Child Development, 91*(2), 382–400. https://doi.org/10.1111/cdev.13175

Benjamin, A. S., & Tullis, J. (2010). What makes distributed practice effective? *Cognitive Psychology, 61*(3), 228–247. https://doi.org/10.1016/j.cogpsych.2010.05.004

Berkowitz, T., Schaeffer, M., Maloney, E., Peterson, L., Gregor, C., Levine, S., & Beilock, S. (2015). Math at home adds up to achievement in school. *Science, 350*(6257), 196–198. https://doi.org/10.1126/science.aac7427

Boaler, J. (2014). Research suggests that timed tests cause math anxiety. *Teaching Children Mathematics, 20*(8), 469–474. https://doi.org/10.5951/teacchilmath.20.8.0469

Bransford, J. D., Brown, A. L., & Cocking, R. R. (2001). *How people learn: Brain, mind, experience, and school* (Expanded ed.). National Academy Press.

Bullock, E. P., Ashby, M. J., Spencer, B., Manderino, K., & Myers, K. (2015). Saxon math in the middle grades: A content analysis. *International Journal of Learning, Teaching and Educational Research, 14*(1), 63–69. https://www.ijlter.org/index.php/ijlter/article/view/472

Busch, J., Schmidt, C., & Grube, D. (2015). Arithmetic fact retrieval: Are there differences between children with developmental dyscalculia and those with mathematical difficulties? *Zeitschrift für Psychologie, 223*(2), 110–119. https://doi.org/10.1027/2151-2604/a000209

Buwalda, T., Borst, J., van der Maas, H., & Taatgen, N. (2016). Explaining mistakes in single digit multiplication: A cognitive model. In D. Reitter & F. E. Ritter (Eds.), *Proceedings of the 14th International Conference on Cognitive Modeling (ICCM 2016)* (pp. 131-136). Penn State.

Byrne, K. A., Silasi-Mansat, C. D., & Worthy, D. A. (2015). Who chokes under pressure? The Big Five personality traits and decision-making under pressure. *Personality and Individual Differences, 74*, 22–28. https://doi.org/10.1016/j.paid.2014.10.009

Callan, M. J., Kim, H., Gheorghiu, A. I., & Matthews, W. J. (2017). The interrelations between social class, personal relative deprivation, and prosociality. *Social Psychological and Personality Science, 8*(6), 660–669. https://doi.org/10.1177/1948550616673877

Cannon, J., & Ginsburg, H. P. (2008). "Doing the math": Maternal beliefs about early mathematics versus language learning. *Early Education and Development, 19*(2), 238–260. https://doi.org/10.1080/10409280801963913

Carpenter, S. K., & Agarwal, P. K. (2020). *How to use spaced retrieval practice to boost learning* (Retrieval Practice). Iowa State University. http://pdf.retrievalpractice.org/SpacingGuide.pdf

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

Chesney, D. L., McNeil, N. M., Matthews, P. G., Byrd, C. E., Petersen, L. A., Wheeler, M. C., Fyfe, E. R., & Dunwiddie, A. E. (2014). Organization matters: Mental organization of addition knowledge relates to understanding math equivalence in symbolic form. *Cognitive Development, 30*, 30–46. https://doi.org/10.1016/j.cogdev.2014.01.001

Cohen, D. A., Pascual-Leone, A., Press, D. Z., & Robertson, E. M. (2005). Off-line learning of motor skill memory: A double dissociation of goal and movement. *Proceedings of the National Academy of Sciences of the United States, 102*(50), 18237–18241. https://doi.org/10.1073/pnas.0506072102

Davoli, C. C., O'Rear, C. D., McAuley, E., McNeil, N. M., & Brockmole, J. R. (2020). Hand position affects performance on multiplication tasks. *Journal of Numerical Cognition, 6*(1), 1–21. https://doi.org/10.5964/jnc.v6i1.211

De Smedt, B., Janssen, R., Bouwens, K., Verschaffel, L., Boets, B., & Ghesquiere, P. (2009). Working memory and individual differences in mathematics achievement: A longitudinal study from first grade to second grade. *Journal of Experimental Child Psychology, 103*(2), 186–201. https://doi.org/10.1016/j.jecp.2009.01.004

Dixon, J. A., & Bangert, A. S. (2005). From regularities to concepts: The development of children's understanding of a mathematical relation. *Cognitive Development, 20*(1), 65–86. https://doi.org/10.1016/j.cogdev.2004.12.001

Donovan, J. J., & Radosevich, D. J. (1999). A meta-analytic review of the distribution of practise effect: Now you see it, now you don't. *Journal of Applied Psychology, 84*(5), 795–805. https://doi.org/10.1037/0021-9010.84.5.795

Dudai, Y., Karni, A., & Born, J. (2015). The consolidation and transformation of memory. *Neuron, 88*(1), 20–32. https://doi.org/10.1016/j.neuron.2015.09.004

Duhon, G. J., Poncy, B. C., Krawiec, C. F., Davis, R. E., Ellis-Hervey, N., & Skinner, C. H. (2022). Toward a more comprehensive evaluation of interventions: A dose-response curve analysis of an explicit timing intervention. *School Psychology Review, 51*(1), 84–94. https://doi.org/10.1080/2372966X.2020.1789435

Ebbinghaus, H. (1964). *Memory: A contribution to experimental psychology* (H. A. Ruger, C. E. Bussenius, & E. R. Hilgard, Trans.). Dover Publications. (Original work published in 1885)

Engle, R. W. (2002). Working memory capacity as executive attention. *Current Directions in Psychological Science, 11*(1), 19–23. https://doi.org/10.1111/1467-8721.00160

Fisher, J. S., & Radvansky, G. A. (2018). Patterns of forgetting. *Journal of Memory and Language, 102*, 130–141. https://doi.org/10.1016/j.jml.2018.05.008

Fisher, J. S., & Radvansky, G. A. (2019). Linear forgetting. *Journal of Memory and Language, 108*, Article 104035. https://doi.org/10.1016/j.jml.2019.104035

Fisher, J. S., & Radvansky, G. A. (2022). Degree of learning and linear forgetting. *Quarterly Journal of Experimental Psychology, 75*(8), 1483–1496. https://doi.org/10.1177/17470218211056464

Fuchs, L. S., Geary, D. C., Compton, D. L., Fuchs, D., Schatschneider, C., Hamlett, C. L., DeSelms, J., Seethaler, P. M., Wilson, J., Craddock, C. F., Bryant, J. D., Luther, K., & Changas, P. (2013). Effects of first-grade number knowledge tutoring with contrasting forms of practice. *Journal of Educational Psychology, 105*(1), 58–77. https://doi.org/10.1037/a0030127

Fuchs, L. S., Wang, A. Y., Preacher, K. J., Malone, A. S., Fuchs, D., & Pachmayr, R. (2021). Addressing challenging mathematics standards with at-risk learners: A randomized controlled trial on the effects of fractions intervention at third grade. *Exceptional Children, 87*(2), 163–182. https://doi.org/10.1177/0014402920924846

Gilmore, C. K. (2006). Investigating children's understanding of inversion using the missing number paradigm. *Cognitive Development, 21*(3), 301–316. https://doi.org/10.1016/j.cogdev.2006.03.007

Gilmore, C. K., Keeble, S., Richardson, S., & Cragg, L. (2017). The interaction of procedural skill, conceptual understanding and working memory in early mathematics achievement. *Journal of Numerical Cognition, 3*(2), 400–416. https://doi.org/10.5964/jnc.v3i2.51

Glenberg, A. M. (1976). Monotonic and nonmonotonic lag effects in paired-associate and recognition memory paradigms. *Journal of Verbal Learning and Verbal Behavior, 15*(1), 1–16. https://doi.org/10.1016/S0022-5371(76)90002-5

Glenberg, A. M. (1979). Component-levels theory of the effects of spacing of repetitions on recall and recognition. *Memory & Cognition, 7*(2), 95–112. https://doi.org/10.3758/BF03197590

Goossens, N. A. M. C., Camp, G., Verkoeijen, P. P. J. L., Tabbers, H. K., Bouwmeester, S., & Zwaan, R. A. (2016). Distributed practice and retrieval practice in primary school vocabulary learning: A multi-classroom study. *Applied Cognitive Psychology, 30*(5), 700–712. https://doi.org/10.1002/acp.3245

Haskell, R. E. (2001). *Transfer of learning: Cognition, instruction, and reasoning.* Academic Press.

Haverty, L. A. (1999). *The importance of basic number knowledge to advanced mathematical problem solving* [Unpublished doctoral dissertation]. Carnegie Mellon University.

Hecht, S. A. (2006). Group differences in adult simple arithmetic: Good retrievers, not-so-good retrievers, and perfectionists. *Memory & Cognition, 34*(1), 207–216. https://doi.org/10.3758/BF03193399

Herbert, D. M. B., & Burt, J. S. (2004). What do students remember? Episodic memory and the development of schematization. *Applied Cognitive Psychology, 18*(1), 77–88. https://doi.org/10.1002/acp.947

Hintzman, D. L. (1974). Theoretical implications of the spacing effect. In R. L. Solso (Ed.), *Theories in Cognitive Psychology: The Loyola Symposium* (pp. 77-99). Lawrence Erlbaum.

Hintzman, D. L., Block, R. A., Summers, J. J. (1973). Modality tags and memory for repetitions: Locus of the spacing effect. *Journal of Verbal Learning and Verbal Behavior, 12*(2), 229–238. https://doi.org/10.1016/S0022-5371(73)80013-1

Horner, A. J., & Doeller, C. F. (2017). Plasticity of hippocampal memories in humans. *Current Opinion in Neurobiology, 43*, 102–109. https://doi.org/10.1016/j.conb.2017.02.004

Imbo, I., Vandierendonck, A., & Fias, W. (2011). Passive hand movements disrupt adults' counting strategies. *Frontiers in Psychology, 2*, Article 201. https://doi.org/10.3389/fpsyg.2011.00201

Kang, S. H. K., Fiske, S. T., Levine, F., Mayer, R., Murphy, K., Newcombe, N., & Worrell, F. (2016). Spaced repetition promotes efficient and effective learning: Policy implications for instruction. *Policy Insights from the Behavioral and Brain Sciences, 3*(1), 12–19. https://doi.org/10.1177/2372732215624708

Kim, A. S. N., Wong-Kee-You, A. M. B., Wiseheart, M., & Rosenbaum, R. S. (2019). The spacing effect stands up to big data. *Behavior Research Methods, 51*, 1485–1497. https://doi.org/10.3758/s13428-018-1184-7

Knowles, N. (2010). *The relationship between timed drill practice and the increase of automaticity of basic multiplication facts for regular education sixth graders* [Doctoral dissertation, Walden University]. Walden Dissertations and Doctoral Studies, 825. https://scholarworks.waldenu.edu/dissertations/825

Korman, M., Raz, N., Flash, T., & Karni, A. (2003). Multiple shifts in the representation of a motor sequence during the acquisition of skilled performance. *Proceedings of the National Academy of Sciences of the United States, 100*(21), 12492–12497. https://doi.org/10.1073/pnas.2035019100

Küpper-Tetzel, C. E., Kapler, I. V., & Wiseheart, M. (2014). Contracting, equal, and expanding learning schedules: The optimal distribution of learning sessions depends on retention interval. *Memory & Cognition, 42*(5), 729–741. https://doi.org/10.3758/s13421-014-0394-1

Kwon, Y. H., Kwon, J. W., & Lee, M. H. (2015). Effectiveness of motor sequential learning according to practice schedules in healthy adults; distributed practice versus massed practice. *Journal of Physical Therapy Science, 27*(3), 769–772. https://doi.org/10.1589/jpts.27.769

LeFevre, J. A., Skwarchuk, S. L., Smith-Chant, B. L., Fast, L., Kamawar, D., & Bisanz, J. (2009). Home numeracy experiences and children's math performance in the early school years. *Canadian Journal of Behavioural Science, 41*(2), 55–66. https://doi.org/10.1037/a0014532

Lynch, G., Kessler, M., Arai, A., & Larson, J. (1990). The nature and causes of hippocampal long-term potentiation. *Progress in Brain Research, 83*, 233–250. https://doi.org/10.1016/S0079-6123(08)61253-4

McNeil, N. M., Fyfe, E. R., & Dunwiddie, A. E. (2015). Arithmetic practice can be modified to promote understanding of math equivalence. *Journal of Educational Psychology, 107*(2), 423–436. https://doi.org/10.1037/a0037687

Metcalfe, J., & Xu, J. (2016). People mind wander more during massed than spaced inductive learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 42*(6), 978–984. https://doi.org/10.1037/xlm0000216

Mozer, M. C., Pashler, H., Cepeda, N., Lindsey, R., & Vul, E. (2009). Predicting the optimal spacing of study: A multiscale context model of memory. In Y. Bengio, D. Schuurmans, J. Lafferty, C. K. I. Williams, & A. Culotta (Eds.), *Advances in Neural Information Processing Systems* (Vol. 22, pp. 1321–1329). Curran Associates.

Price, G. R., Mazzocco, M. M., & Ansari, D. (2013). Why mental arithmetic counts: Brain activation during single digit arithmetic predicts high school math scores. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience, 33*(1), 156–163. https://doi.org/10.1523/JNEUROSCI.2936-12.2013

Radvansky, G. A., Doolen, A. C., Pettijohn, K. A., & Ritchey, M. (2022). A new look at memory retention and forgetting. *Journal of Experimental Psychology: Learning, Memory, and Cognition. 48*(11), 1698–1723. https://doi.org/10.1037/xlm0001110

Raghubar, K. P., Barnes, M. A., & Hecht, S. A. (2010). Working memory and mathematics: A review of developmental, individual difference, and cognitive approaches. *Learning and Individual Differences, 20*(2), 110–122. https://doi.org/10.1016/j.lindif.2009.10.005

Rea, C. P., & Modigliani, V. (1985). The effect of expanded versus massed practice on the retention of multiplication facts and spelling lists. *Human Learning, 4*(1), 11–18.

Rickard, T. C. (1997). Bending the power law: A CMPL theory of strategy shifts and the automatization of cognitive skills. *Journal of Experimental Psychology: General, 126*(3), 288–311. https://doi.org/10.1037/0096-3445.126.3.288

Rickard, T. C., Lau, J. S.-H., & Pashler, H. (2008). Spacing and the transition from calculation to retrieval. *Psychonomic Bulletin & Review, 15*(3), 656–661. https://doi.org/10.3758/PBR.15.3.656

Roediger, H. L., III, & Karpicke, J. D. (2006). Test-enhanced learning: Taking memory tests improves long-term retention. *Psychological Science, 17*(3), 249–255. https://doi.org/10.1111/j.1467-9280.2006.01693.x

Rohrer, D., & Taylor, K. (2006). The effects of overlearning and distributed practise on the retention of mathematics knowledge. *Applied Cognitive Psychology, 20*(9), 1209–1224. https://doi.org/10.1002/acp.1266

Scoville, W. B., & Milner, B. (1957). Loss of recent memory after bilateral hippocampal lesions. *Journal of Neurology, Neurosurgery, and Psychiatry, 20*(1), 11–21. https://doi.org/10.1136/jnnp.20.1.11

Shaughnessy, J. J., Zimmerman, J., & Underwood, B. J. (1972). Further evidence on the MP-DP effect in free-recall learning. *Journal of Verbal Learning and Verbal Behavior, 11*(1), 1–12. https://doi.org/10.1016/S0022-5371(72)80053-7

Siegler, R. S. (1988). Individual differences in strategy choices: Good students, not-so-good students, and perfectionists. *Child Development, 59*(4), 833–851. https://doi.org/10.2307/1130252

Spielberger, C. D., Gorsuch, R. L., Lushene, R., Vagg, P. R., & Jacobs, G. A. (1983). *Manual for the State-Trait Anxiety Inventory*. Consulting Psychologists Press.

Squire, L. R., Genzel, L., Wixted, J. T., & Morris, R. G. (2015). Memory consolidation. *Cold Spring Harbor Perspectives in Biology, 7*(8), Article a021766. https://doi.org/10.1101/cshperspect.a021766

Sweegers, C. C. G., Takashima, A., Fernández, G., & Talamini, L. M. (2014). Neural mechanisms supporting the extraction of general knowledge across episodic memories. *NeuroImage, 87*, 138–146. https://doi.org/10.1016/j.neuroimage.2013.10.063

Tluczek, A., Henriques, J. B., & Brown, R. L. (2009). Support for the reliability and validity of a six-item state anxiety scale derived from the State-Trait Anxiety Inventory. *Journal of Nursing Measurement, 17*(1), 19–28. https://doi.org/10.1891/1061-3749.17.1.19

Toppino, T. C., & Bloom, L. C. (2002). The spacing effect, free recall, and two-process theory: A closer look. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 28*(3), 437–444. https://doi.org/10.1037/0278-7393.28.3.437

Toppino, T. C., Phelan, H.-A., & Gerbier, E. (2018). Level of initial training moderates the effects of distributing practice over multiple days with expanding, contracting, and uniform schedules: Evidence for study-phase retrieval. *Memory and Cognition, 46*(6), 969–978. https://doi.org/10.3758/s13421-018-0815-7

Tronsky, L. N., & Royer, J. M. (2003). Relationships among basic computational automaticity, working memory, and complex mathematical problem solving: What we know and what we need to know. In J. M. Royer (Ed.), *Mathematical cognition* (pp. 117–146). Information Age Publishing.

Tulving, E. (1972). *Episodic and semantic memory.* In E. Tulving & W. Donaldson (Eds.), *Organization of memory* (pp. 381–403). Academic Press.

Verkoeijen, P. P. J. L., Rikers, R. M. J. P., & Özsoy, B. (2008). Distributed rereading can hurt the spacing effect in text memory. *Applied Cognitive Psychology, 22*(5), 685–695. https://doi.org/10.1002/acp.1388

Wang, Z., Lukowski, S. L., Hart, S. A., Lyons, I. M., Thompson, L. A., Kovas, Y., Mazzocco, M. M. M., Plomin, R., & Petrill, S. A. (2015). Is math anxiety always bad for math learning? The role of math motivation. *Psychological Science, 26*(12), 1863–1876. https://doi.org/10.1177/0956797615602471

Widmer, P. S., Semmer, N. K., Kälin, W., Jacobshagen, N., & Meier, L. L. (2012). The ambivalence of challenge stressors: Time pressure associated with both negative and positive well-being. *Journal of Vocational Behavior, 80*(2), 422–433. https://doi.org/10.1016/j.jvb.2011.09.006

Wiseheart, M., Küpper-Tetzel, C. E., Weston, T., Kim, A. S. N., Kapler, I. V., & Foot-Seymour, V. (2019). Enhancing the quality of student learning using distributed practice. In J. Dunlosky & K. Rawson (Eds.), *Cambridge handbook of cognition and education* (pp. 550–584). Cambridge University Press.

Yerkes, R. M., & Dodson, J. D. (1908). The relation of strength of stimulus to rapidity of habit formation. *Journal of Comparative Neurology and Psychology, 18*(5), 459–482. https://doi.org/10.1002/cne.920180503

Young, D. L., Goodie, A. S., Hall, D. B., & Wu, E. (2012). Decision making under time pressure, modeled in a prospect theory framework. *Organizational Behavior and Human Decision Processes, 118*(2), 179–188. https://doi.org/10.1016/j.obhdp.2012.03.005

PsychOpen GOLD

Appendix

Table A1

Participant Characteristics With Means (SD) or Percentage for Each Timing Grouping Variable for 1-, 2- and 7-Day Spacing Conditions

	Timed						Untimed					
	1 Day		2 Day		7 Day		1 Day		2 Day		7 Day	
Variable	M (SD)	%	M (SD)	%	M (SD)	%	M (SD)	%	M (SD)	%	M (SD)	%
Age (Years)	24.5 (5.1)		26.8 (6.8)		23.4 (4.6)		25.8 (5.8)		22.8 (4.3)		25.2 (4.9)	
Racial/Ethnic Makeup												
American Indian or Alaskan Native		0		0		0		0		0		0
Asian		8		4		12		0		0		8
Black or African American		4		0		4		0		4		0
Hispanic or Latino		16		19		8		40		8		12
Prefer Not to Answer		0		0		4		0		0		0
White		72		77		72		60		89		80
Interest in Mathematics	2.6 (1.0)		2.3 (1.5)		2.1 (1.2)		2.5 (1.0)		2.4 (1.3)		2.2 (1.2)	
Math Anxiety	2.0 (1.1)		2.0 (1.2)		2.2 (1.2)		1.9 (1.0)		1.9 (1.2)		2.0 (1.1)	
Level of Education	2.9 (1.3)		2.5 (1.8)		2.6 (1.5)		3.4 (1.6)		2.4 (1.6)		2.8 (1.5)	
Socioeconomic Status (Parents' Education)	3.2 (2.0)		2.3 (1.4)		2.8 (2.0)		3.0 (1.9)		2.3 (2.1)		2.2 (1.8)	
Household Income	2.2 (1.5)		2.5 (1.3)		2.2 (1.6)		3.3 (2.0)		2.4 (1.5)		2.3 (1.5)	
Occupation												
Agriculture		0		0		4		0		0		0
Arts, Culture, and Entertainment		0		8		8		12		4		4
Business, Management, and Administration		8		15		4		36		4		8
Communications		0		0		8		0		4		8
Community and Social Services		0		4		4		0		8		4
Education		13		12		25		16		15		16
Engineering		33		27		13		16		27		24
Government		0		4		0		0		0		0
Health and Medicine		8		12		0		12		12		8
Homemaker		4		0		4		0		0		0
Installation, Repair, and Maintenance		4		8		0		0		0		4
Sales		8		0		4		4		8		8
Student		21		8		17		4		12		16
Unemployed		0		0		0		0		4		0
Unspecified Other		0		4		8		0		4		0
Use of Mathematics in Current Occupation	2.2 (1.1)		1.5 (1.2)		1.7 (1.1)		2.0 (1.2)		2.2 (1.2)		1.9 (1.0)	
Current Place of Residence												
Africa		4		0		0		4		0		0
Europe		76		81		100		64		96		84
North America		16		15		0		24		4		12
South America		4		4		0		8		0		4

Note. **Interest in Mathematics** was in the range of 0-4 (0 = very uninterested, 1 = somewhat uninterested, 2 = neutral, 3 = somewhat interested, 4 = very interested). **Math Anxiety** was in the range of 0-4 (0 = very confident, 1 = somewhat confident, 2 = neutral, 3 = somewhat anxious, 4 = very anxious). **Education** was in the range of 0-6 (0 = less than a high school diploma; 1 = high school degree or equivalent (e.g., GED); 2 = some college, no degree; 3 = associate degree (e.g., AA, AS); 4 = Bachelor's degree (e.g., BA, BS); 5 = master's degree (e.g., MA, MS, Med); 6 = doctorate or professional degree (e.g., MD, DDS, PhD)). **Household Income** was in the range of 1-7 (1 ≤ $20,000; 2 = $20,000-34,999; 3 = $35,000-49,999; 4 = $50,000-74,999; 5 = $75,000-99,999; 6 = $100,000-149,999; 7 ≥ $150,000), 17 of all participants preferred not to answer.

Table A2

Participant Characteristics With Means (SD) or Percentage for Each Timing Grouping Variable for 10- and 14-Day Spacing Conditions

	Timed				Untimed			
	10 Day		14 Day		10 Day		14 Day	
Variable	*M (SD)*	%	*M (SD)*	%	*M (SD)*	%	*M (SD)*	%
Age (Years)	23.5 (5.8)		26.9 (7.3)		26.8 (6.5)		26.1 (6.8)	
Racial/Ethnic Makeup								
American Indian or Alaskan Native		0		4		0		0
Asian		8		4		0		0
Black or African American		13		0		8		0
Hispanic or Latino		13		12		8		8
Prefer Not to Answer		0		0		0		12
White		67		80		83		80
Interest in Mathematics	2.5 (1.3)		2.0 (1.3)		2.0 (1.2)		2.1 (1.3)	
Math Anxiety	2.0 (1.0)		2.1 (1.3)		2.0 (1.1)		2.2 (1.1)	
Level of Education	2.4 (1.3)		3.3 (1.6)		2.5 (1.6)		2.2 (1.5)	
Socioeconomic Status (Parents' Education)	2.6 (2.2)		2.6 (1.6)		1.9 (1.4)		2.5 (2.0)	
Household Income	2.9 (1.7)		2.3 (1.2)		2.4 (1.3)		2.6 (1.6)	
Occupation								
Agriculture		4		0		4		4
Arts, Culture, and Entertainment		9		8		4		0
Business, Management, and Administration		9		8		17		12
Communications		13		0		8		4
Community and Social Services		0		4		8		4
Education		17		32		21		4
Engineering		9		12		8		20
Government		0		0		4		0
Health and Medicine		13		8		8		4
Homemaker		0		8		4		8
Installation, Repair, and Maintenance		0		0		0		8
Sales		4		8		8		12
Student		22		12		4		12
Unemployed		0		0		0		4
Unspecified Other		0		0		0		4
Use of Mathematics in Current Occupation	1.9 (1.5)		1.8 (0.9)		2.0 (1.3)		1.8 (1.3)	
Current Place of Residence								
Africa		0		0		0		0
Europe		83		92		92		92
North America		13		8		8		8
South America		4		0		0		0

Note. **Interest in Mathematics** was in the range of 0-4 (0 = very uninterested, 1 = somewhat uninterested, 2 = neutral, 3 = somewhat interested, 4 = very interested). **Math Anxiety** was in the range of 0-4 (0 = very confident, 1 = somewhat confident, 2 = neutral, 3 = somewhat anxious, 4 = very anxious). **Education** was in the range of 0-6 (0 = less than a high school diploma; 1 = high school degree or equivalent (e.g., GED); 2 = some college, no degree; 3 = associate degree (e.g., AA, AS); 4 = Bachelor's degree (e.g., BA, BS); 5 = master's degree (e.g., MA, MS, Med); 6 = doctorate or professional degree (e.g., MD, DDS, PhD)). **Household Income** was in the range of 1-7 (1 ≤ \$20,000; 2 = \$20,000-34,999; 3 = \$35,000-49,999; 4 = \$50,000-74,999; 5 = \$75,000-99,999; 6 = \$100,000-149,999; 7 ≥ \$150,000), 17 of all participants preferred not to answer.

MCLS

Journal of Numerical Cognition (JNC) is an official journal of the Mathematical Cognition and Learning Society (MCLS).

leibniz-psychology.org

PsychOpen GOLD is a publishing service by Leibniz Institute for Psychology (ZPID), Germany.