Goal-Driven Regulation of Learning: How Goal Orientation Shapes Time Allocation between Retrieval and Study Practice

A dissertation submitted by

Lan Anh Ly Do

in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

in

Psychology

Tufts University

May 2025

Advisor: Dr. Ayanna K. Thomas

Abstract

Effortful study strategies, such as retrieval practice, can enhance long-term retention of studied material. However, it may lead to short-term setbacks and immediate failures, which can evoke emotional discomfort. As a result, learners may be reluctant to engage in retrieval practice, despite the long-term benefits for learning. The goals of successful learning and emotional well-being may sometimes compete, influencing the strategies learners choose to use. This dissertation examined how learners allocated study time to different strategies under varying goal contexts. Specifically, I explored how learners divided time between retrieval practice and an alternative, less emotionally taxing strategy, referred to as study practice, across four experiments. Experiments 1 through 3 manipulated goal orientation (learning or well-being) in a vocabulary learning task. I predicted that participants encouraged to safeguard emotional well-being would allocate less time to retrieval practice than those encouraged to maximize learning. The results supported this prediction. This pattern held across several variations in design: when time constraints were varied (Experiment 1A) or identical (1B), when participants received information about the benefits and costs of retrieval (2), and whether participants allocated time for a hypothetical student (1A, 1B, 2) or for themselves (3). Across Experiments 1-3, most participants viewed retrieval practice as more effective for learning but also more likely to cause anxiety and frustration. Experiment 4 extended this investigation to a university course setting. Participants reported their learning and well-being goals at the beginning of each class, and chose between interval testing (retrieval practice) and interval study (study practice). Results showed that well-being goals remained stable throughout the semester, but learning goals declined over time. Participants were more likely to choose retrieval practice when well-being goals were lower. Learning goals, in contrast, did not significantly predict strategy choice,

possibly because most participants believed study practice was more effective in this context. In summary, this dissertation shows how cognitive and emotional goals jointly shape study strategy use. The findings support models of motivated self-regulated learning and underscore the importance of addressing both the learning benefits and emotional costs of effortful strategies to help learners persist with them over time.

Acknowledgements

I am deeply grateful to my advisor, Dr. Ayanna K. Thomas, for her steadfast support, thoughtful guidance, and unwavering belief in me throughout this journey. Her groundbreaking work on the integration of cognitive and emotional aspects of learning laid the foundation for this dissertation. Her mentorship has profoundly shaped both my research and the kind of scholar I strive to become. What I've learned from her extends beyond research—she leads with generosity, especially in moments of uncertainty or mistakes, always allowing space for growth. I hope to carry forward that same spirit in my own mentorship: to be patient, encouraging, and open-minded, and to support others with the same care and trust she has shown me.

I also wish to sincerely thank my committee members—Dr. Holly Taylor, Dr. Heather Urry, and Dr. Veronica Yan—for their generous feedback, insightful questions, and steady encouragement throughout this process. Their expertise and thoughtful perspectives challenged me to think more critically, consider alternative interpretations, and refine the ideas in this dissertation. I'm grateful for their time, care, and investment in helping me grow as a researcher.

This dissertation is not only the product of academic inquiry, but also the culmination of a long personal journey shaped by the many students I've had the privilege to teach and mentor. Working with them fueled my passion for understanding how people learn—and more importantly, how they *feel* while learning. They inspired the questions that drove this research, especially those surrounding why we sometimes avoid the very strategies that help us learn best. Their curiosity, struggles, and determination gave this work purpose.

I am also incredibly thankful to my labmates—Dr. Alia Wulff, Dr. McKinzey Torrance, and Himanshu Chaudhary—for their thoughtful feedback, collaborative spirit, and the many energizing conversations that deeply influenced the direction of this project. Their insights and

encouragement consistently pushed my thinking forward. I'm especially grateful to our exceptional lab manager, Sophia Grekin, for her unwavering support and invaluable assistance in developing and refining the experimental materials. I would also like to extend my sincere thanks to the dedicated research assistants who generously contributed their time and care to piloting studies, reviewing data, and supporting the research process at every stage. Their contributions were essential to bringing this work to life.

I owe so much to my family and my partner for their unwavering love and support. Their belief in me never wavered, even during the most difficult and uncertain moments. They stood by me through every high and low, offering steady encouragement, listening patiently, and reminding me of my strength when I needed it most. Their warmth and presence were a constant source of comfort and motivation throughout this journey. I truly could not have reached this point without their enduring care, understanding, and the countless ways they helped me stay grounded and hopeful.

This dissertation may mark the end of a chapter, but the people who brought it to life—my mentors, peers, students, and mentees—and the questions they inspired will continue to shape my work and guide my thinking for years to come.

Table of Contents

List of Tables

List of Figures

Goal-Driven Regulation of Learning: How Goal Orientation Shapes Time Allocation between Retrieval and Study Practice

Research in cognitive and educational psychology has identified a range of effective study strategies that promote long-term memory retention of studied material and facilitate the transfer of knowledge to different contexts (Dunlosky et al., 2013). For example, learners can remember information better when they are allowed to actively generate all or part of some to-be-remembered information rather than merely reading that information (*the generation effect*; Jacoby, 1978; Slamecka & Graf, 1978; also see Bertsch et al., 2007). Similar benefits for memory also emerge when one performs self-testing by attempting to retrieve previously studied information instead of restudying it (*the testing effect*; Butler, 2010; Roediger & Karpicke, 2006b). Spacing out study sessions over time rather than cramming or massing them together in one single session increases memorability of information in the long term (*the spacing effect*; Cepeda et al., 2009; Kornell, 2009; McDaniel et al., 2013). Further, intermixing various topics during practice (i.e., interleaved study) rather than grouping or blocking them by topic (i.e., blocked study) can also promote durable memory and deeper understanding of the learning content (*the interleaving effect*; Kornell & Bjork, 2008; Kang & Pashler, 2012).

One common feature across these study strategies (generation, retrieval practice, spacing, and interleaving) is that they make learning more effortful in a beneficial way. By requiring learners to invest greater mental effort during study, these strategies improve how information is encoded and strengthen memory traces. As a result, they lead to more durable learning and better long-term retention of the studied material—a phenomenon known as *desirable difficulty* (Bjork, 1975; Bjork & Bjork, 2011).

However, long-term benefits of effortful and challenging learning often come at the expense of short-term performance and learner experience. Using effortful study strategies may slow down the rate of acquisition and impede initial performance during training (Bjork & Linn, 2006; Yan & Schuetze, 2022). Much empirical research has demonstrated that effortful learning strategies lead to a higher perception of cognitive demands (e.g., Kirk-Johnson et al., 2019; Yan & Schuetze, 2022) and can elicit negative emotions (see Wenzel, 2022 for an overview). The present study examined how learning and well-being goals influenced the way learners navigated the trade-off between learning benefits and negative emotional experiences when regulating the use of retrieval practice, an effective yet demanding study strategy that can induce anxiety and frustration.

Retrieval Practice: The Trade-Off Between Learning and Well-Being

Research consistently demonstrates that students retain information more effectively when they actively retrieve it from memory, such as through self-testing, compared to simply rereading the material for an equivalent amount of time (see Roediger et al., 2011 for a review). One key reason for this robust mnemonic benefit is that retrieval requires greater cognitive effort than restudying. The more effort involved in recalling information, the more deeply learners engage with the material during the retrieval attempt. This effortful process prompts more elaborate cognitive processing and strengthens the memory trace, making learning more durable over time (Pyc & Rawson, 2009). Retrieval tends to be especially effective when it is desirably difficult. For example, retrieving information after a longer delay results in better long-term retention than retrieving it shortly after study, when the material is still easily accessible (Pyc & Rawson, 2009). This pattern reflects the broader principle that retrieval is most beneficial when it challenges the learner—an idea emphasized by Bjork and Bjork (2011), who noted that retrieval

practice enhances storage strength (the durability of memory) most when retrieval strength (its immediate accessibility) is low.

The benefits of retrieval practice are evident even after a single study-test session (Carpenter, 2009; Carpenter & DeLosh, 2006) and become more pronounced when retrieval is repeated across multiple trials (Karpicke & Roediger, 2008). While retrieval is effective even without feedback (Butler & Roediger, 2008), providing feedback, such as supplying the correct answer, further enhances learning by helping students correct errors and reinforce accurate information (Bangert-Drowns et al., 1991; Kulhavy & Stock, 1989; Pashler et al., 2005). In addition to improving retention, retrieval practice also facilitates both near and far transfer of knowledge, allowing learners to better apply what they have learned to new inference problems within and across domains (Barnett & Ceci, 2002; Butler, 2010).

The benefits of retrieval practice extend beyond the laboratory. Classroom studies using authentic course materials and assessments have shown that retrieval practice improves academic performance (Bangert-Drowns, 1991; Lyle & Crawford, 2011; McDaniel et al., 2007; Thomas et al., 2020). Students who take regular quizzes tend to perform better than those who do not (e.g., Lyle & Crawford, 2011). Moreover, compared to having only one or two major exams (e.g., a midterm and a final), frequent testing promotes sustained engagement by encouraging regular study and fostering more consistent study habits throughout the semester (Mawhinney et al., 1971; Michael, 1991). Taken together, this body of evidence highlights retrieval practice as not only a means of assessment but also a highly effective strategy for enhancing long-term learning and academic success.

However, the demanding and challenging nature of retrieval, while making it effective for learning, may also have emotional costs (Wenzel, 2022). Compared to passive strategies like

rereading or copying content, retrieving information from memory more often results in learner mistakes and reveals gaps in knowledge. Although these experiences can enhance metacognitive awareness by helping learners recognize what they do and do not know (Koriat & Bjork, 2006), they may also lower confidence (Miller & Geraci, 2014) and trigger negative emotions such as frustration or anxiety (Hinze & Rapp, 2014). Taking practice tests, while beneficial for reinforcing learning, can also be a source of anxiety and stress (Ott, 2017; Sarason & Sarason, 1990). Studies have shown that both high- and low-stakes tests can evoke negative emotions such as anger, uncertainty, annoyance, and pressure (Wenzel & Reinhard, 2021a, 2021b). Research further suggests that students with higher test anxiety tend to use self-testing less often (Liu et al., 2024). Moreover, these emotional challenges can be intensified by time constraints. Time pressure during practice tests or exams can intensify feelings of anxiety and stress, resulting in poorer performance (Caviola et al., 2017; Orfus, 2008). Therefore, while retrieval practice is an effective learning strategy, its emotional strain may influence learners' willingness to use it, even when they are aware of its benefits for learning.

Highman et al. (2022) demonstrated the mnemonic benefits of retrieval practice in real classroom setting, showing that repeated retrieval practice over three sessions led to better recall of course material at the end of the semester compared to repeated restudying. However, the study also highlighted a trade-off between learning and well-being, particularly during the early stages of learning. Learners reported higher anxiety and lower feelings of mastery in session 1 when engaging in retrieval practice, compared to repeated studying. These negative effects on anxiety and mastery diminished in later sessions (2 and 3), suggesting that the discomfort of retrieval practice may be temporary, with benefits emerging over time.

These findings are significant as they raise the question of whether learners would choose retrieval practice if given the option to select their study strategy, rather than being required to use it. Would they recognize its benefits for learning and choose it to optimize their academic outcomes? Even if they acknowledge its long-term advantages, would the initial experience of heightened anxiety and reduced feelings of mastery cause some learners to hesitate in adopting retrieval practice? To address these questions, the following section reviews the literature on learners' perceptions of retrieval practice and explores how learners make strategic choices in light of the trade-off between its impact on memory and emotional experiences.

Learners' Perceptions and Use of Retrieval Practice

The literature on learners' awareness of the benefits of retrieval practice is mixed. Early research indicated that students often failed to recognize the advantages of testing. For instance, in studies by Agarwal et al. (2008) and Roediger and Karpicke (2006a), participants who reread a text passage multiple times believed they would remember it better than participants who read the passage and then took one or more tests. This suggests that students underestimated the benefits of retrieval practice when engaging in reading comprehension tasks. McCabe (2011) similarly found that most students perceived restudying as more effective than retrieval practice for retention, further highlighting a lack of awareness of testing's advantages. However, McCabe also demonstrated that awareness increased after students received targeted instruction on cognitive psychology principles, such as the testing effect. Students who participated in these lectures showed significantly higher recognition of retrieval practice's benefits, suggesting that awareness improves with explicit instruction.

More recent studies indicate that many learners now recognize the benefits of retrieval practice and actively use it to enhance their learning. For example, using the same scenarios from

McCabe (2011), Morehead et al. (2016) found that most instructors (65%) and students (72%) endorsed the value of testing as a study strategy. Kuhbandner and Emmerdinger (2019) found that students preferred rereading early in the learning process, but most switched to retrieval practice later, rereading only the parts they did not understand. Other studies examining students' actual study behaviors (e.g., Bottiroli et al., 2010; Janes et al., 2018; Kornell & Son, 2009) also suggest that students use a considerable amount of self-testing. Kornell and Son (2009), for instance, found that participants used retrieval practice in over 50% of study trials (55%) in a vocabulary learning task, highlighting its appeal as an effective learning strategy. However, it is important to note that when asked why they engaged in self-testing, most students reported using it as a metacognitive strategy to track their learning progress (Hartwig & Dunlosky, 2012; Kornell & Bjork, 2007; Kornell & Son, 2009), indicating that they were not fully aware of its benefits for memory beyond its role as an evaluation tool.

Research demonstrates that learners' perceptions and use of retrieval practice depend on a variety of factors. For example, Son (2005) showed that students used retrieval practice selectively, preferring to test themselves on items they felt confident about and to re-study items they struggled with. This indicates that learners tend to engage in retrieval practice primarily when they anticipate success, highlighting that their use of retrieval practice is influenced by their confidence in their ability to retrieve the information. Additionally, factors such as item difficulty (Toppino et al., 2018) and perceived costs, such as time, effort, and emotional responses, can also affect when and how learners choose to use retrieval practice (Rea et al., 2023).

For instance, Rea et al. (2023) found that students believed high-achieving individuals were more likely to use effortful study strategies, such as self-testing, self-explanation, spaced

study, and spaced retrieval practice, compared to average or lower-achieving students (Experiment 1). This finding was replicated in Experiment 2, where participants acknowledged that active strategies were more effective for learning than passive strategies like reviewing material, highlighting, or underlining. However, many participants reported that their own study behaviors aligned more with passive strategies. This was largely because they believed that effective strategies required more time and effort, and could lead to feelings of anxiety and nervousness.

Wang et al. (2023) developed an intervention to enhance learners' knowledge of the benefits of retrieval practice and teach them how to implement it more effectively, while reducing perceived costs in terms of effort and time. Although the intervention successfully increased learners' intention to use retrieval practice, it did not significantly reduce perceived costs or increase actual use of the strategy. This suggests that while increased awareness may improve learners' intent to use retrieval practice, it is not sufficient to change their actual study behavior.

Together, these studies demonstrate that even when students are aware of the benefits of retrieval practice, they may hesitate to use it due to various associated costs. Emotional costs, such as stress and frustration, can discourage students from engaging in effortful strategies, even when they understand the long-term benefits. With that in mind, the next section discusses how motivation and goals related to learning and well-being can influence learning decisions, especially when students face the conflict between learning benefits and emotional costs.

The Effects of Motivation and Goals on Learning Behaviors

Research indicates that learning and well-being goals, alongside learning experiences, interact dynamically to influence how learners regulate their study strategies. Several theoretical

frameworks of self-regulated learning (SRL) emphasize the importance of integrating cognitive and affective goals in the regulation process. Notable models include the dual-processing self-regulation model (Boekaerts, 2011; Boekaerts & Cascallar, 2006), the metacognitive and affective model of self-regulated learning (MASRL; Efklides, 2011), the meta-affect metacognition framework (MAMC; Thomas et al., 2022), and the multiple goals regulation framework (Kim et al., 2023). These frameworks collectively underscore the critical role that both cognitive and emotional goals play in guiding SRL. For a broader perspective, Panadero (2017) offers a comprehensive review of SRL models, further highlighting the interplay between cognitive and affective factors.

SRL is broadly defined as the process through which individuals take control of their learning by setting goals, monitoring progress, and adjusting strategies to achieve desired outcomes. A key focus of SRL frameworks is the interaction between learning goals, such as mastering material or improving performance, and well-being goals, such as reducing stress or maintaining a positive mood. These goals play a crucial role in guiding decisions about what, when, and how to study (Boekaerts, 2011; Thomas et al., 2022).

A growing body of research highlights the critical role of motivation and goal orientation in shaping learners' cognitive engagement and study strategies (Greene et al., 2004; Phan, 2009; Pintrich & Schunk, 2002). For instance, students are more likely to use effective study strategies in courses they consider important (Rea et al., 2022) and tend to view effort more positively in classes they find motivating (Flake, 2015). Students with a mastery goal orientation—defined by a focus on understanding and mastering the material—are more likely to sustain effort and use effective learning strategies such as rehearsal, elaboration, organization, critical thinking, and metacognitive self-regulation (Kadioglu & Kondakci, 2014). Students with a mastery goal are

also more inclined to seek out challenging tasks and devote more time and effort to their studies (Ames & Archer, 1988; Harackiewicz et al., 1997). Additionally, mastery-oriented students also tend to engage more deeply with the material, actively integrating information from multiple sources (e.g., readings and lectures) and processing it at a deeper level (Pintrich & DeGroot, 1990).

In line with this, Weissgerber et al. (2016) found that students with stronger mastery goals reported more positive attitudes and greater engagement with desirable difficulties. In contrast, students driven by avoidance goals—such as trying to avoid negative evaluations—tended to show lower levels of engagement. Further evidence from Weissgerber et al. (2018) suggests that students with a higher need for cognition not only responded more positively to challenging tasks but were also more likely to adopt effortful and effective learning strategies.

Beyond intrinsic goals like mastery orientation, extrinsic factors can also shape learners' motivation and behavior. For instance, motivation to learn can be enhanced by external incentives such as higher point values, monetary rewards, or an increased likelihood of being tested. Research has shown that when motivation to learn particular material increases, so does the effort students devote to learning it. Specifically, participants allocated more study time to information they were motivated to learn and restudied them more frequently (Ariel et al., 2009; Castel et al., 2013; Halamish et al., 2019; Koriat et al., 2006).

However, higher motivation to learn does not always lead to increased use of retrieval practice, particularly when learners are unaware of its benefits. For example, Liu et al. (2024) manipulated task importance by varying the stakes of a mid-term exam to examine its impact on strategy choice. When reviewing high-stakes topics that contributed 30% to their final grade, students were more likely to choose restudying over self-testing. In contrast, they engaged in

more self-testing when the topics were low-stakes and did not contribute to their final grade. This may be because students viewed self-testing as a tool for diagnosing their learning, while they considered restudying to be more effective than self-testing for improving learning (Badali et al., 2022). As a result, they chose self-testing to assess their knowledge of less important low-stakes topics. Conversely, for high-stakes topics, they likely focused on restudying to enhance performance.

Although SRL theories and frameworks often emphasize the interaction between learning and well-being goals, emotional goals have received comparatively less attention in SRL research. Some learners may hesitate to embrace the challenges of learning or persist through difficulties, even when doing so could support their academic success (Weissgerber et al., 2016). When learners prioritize emotional well-being, anxiety and frustration associated with effortful study strategies, such as retrieval practice, may deter them from using these effective strategies, even if they recognize their benefits (Rea et al., 2022; Wang et al., 2023). This underscores the need for a more holistic approach, in which students balance both learning outcomes and emotional well-being when deciding how to study.

The cognitive and emotional impacts of effortful learning can lead to different learning decisions, depending on learners' motivations related to both learning and well-being. For instance, while retrieval practice may prompt learners to realize that additional study is needed, they may also recognize that more study could lead to increased emotional strain. A learner’s goals and motivations can influence how this tension is resolved. For example, learners with mastery-oriented goals, who aim to deeply understand the material, may value the cognitive benefits of effortful study strategies more than the emotional costs, resulting in greater engagement with effective and effortful strategies (Weissgerber et al., 2016). Conversely,

learners who prioritize emotional well-being may weigh the emotional costs more heavily and, as a result, be less willing to engage in additional effortful study strategies, even when these strategies are more effective (Boekaerts, 2011; Efklides, 2011; Kim et al., 2023).

Building on this insight, the present study contributes to existing literature by addressing a frequently overlooked aspect of effortful learning—emotional costs—and examining how the trade-off between learning and emotional experiences influences the adoption of effortful study strategies. While research suggests that motivation and goals related to both learning and emotional well-being influence students' engagement with effortful study strategies, there is a gap in experimental studies that directly explore this dynamic. This study sought to fill this gap by investigating how cognitive and affective goals influenced the self-regulation of challenging yet effective study strategies, thereby providing a deeper understanding of motivated SRL in the context of effortful learning.

Overview of the Experiments

The present study explored how goals related to learning and emotional well-being affected the allocation of study time between retrieval and a less effective but more enjoyable alternative, study practice, across four experiments. In Experiments 1 through 3, I directly manipulated goals (learning goal vs. well-being goal) to examine the impact of goal orientation on strategy choice. Experiment 4, conducted in an in-person classroom setting, investigated how learning and well-being goals evolved over the course of a semester and how these changes influenced strategy choice between retrieval practice and study practice.

In Experiments 1A and 1B, participants learned a series of Swahili-English word pairs. Half of the pairs were learned through repeated testing (retrieval practice), while the other half were learned through repeated studying (study practice). Following this phase, participants were

instructed to prioritize either learning, emotional well-being, or given no specific goal. Participants then allocated study time between retrieval practice and study practice to help a hypothetical student achieve the designated goal. The only difference between Experiments 1A and 1B was the manipulation of time limits. In Experiment 1A, the time limits differed between the two study strategies to create distinct learning experiences, whereas in Experiment 1B, the time limit was held constant. This design allowed me to provide direct evidence supporting the causal effect of goal orientation on the way learners study.

Drawing on previous research indicating participants' preference for retrieval practice in vocabulary learning tasks (e.g., Kornell & Son, 2009), I predicted that participants would allocate at least 50% of their study time to retrieval practice, even in the absence of a specific goal. Emphasizing a learning goal was expected to further promote the use of retrieval practice. In contrast, prioritizing emotional well-being might lead to hesitation in engaging with retrieval practice due to its demanding nature and the potential to induce frustration and anxiety (Wenzel, 2022). I also expected that varying time limits would influence participants' perceptions of anxiety and strategy effectiveness. Specifically, shorter time limits could increase anxiety, making the emotional costs of retrieval practice more pronounced, while simultaneously reducing performance and diminishing its perceived effectiveness (Onwuegbuzie & Seaman, 2010). By manipulating the time limit, we further examined how goal orientation influenced the regulation of strategy use under different circumstances.

Experiment 2 was designed to isolate the effects of goal orientation by ensuring participants understood how retrieval practice could influence both learning and emotional well-being. Research indicates that many individuals are unaware of the mnemonic benefits of retrieval practice (see Rivers, 2021 for a review). If participants do not recognize these

advantages, they may avoid using retrieval practice, even when encouraged to prioritize learning. Additionally, Experiment 2 controlled for the effect of performance during retrieval practice on its usage, as research has shown that retrieval success can influence strategy choice (Son, 2005; Toppino et al., 2018).

As a result, in Experiment 2, participants were provided with descriptions of the two strategies, alongside research information explaining that retrieval practice is more effective for learning yet more likely to trigger anxiety and frustration. This condition was compared to a control group that did not receive such information. By controlling the information given to participants, Experiment 2 offered clearer insights into how goal orientation influenced strategy regulation, depending on whether participants fully understood the impacts of retrieval practice on learning and well-being. I predicted that goal orientation would still exert its effects regardless of the information provided, but the presence of this information would strengthen the impact of goal orientation.

In Experiment 3, I aimed to extend the findings of Experiments 1A and 1B by modifying the scenario to instruct participants to allocate study time for their own learning rather than for a hypothetical student. The design remained similar to Experiment 1B, with a consistent time limit for both retrieval and study practice. Although prior research has shown that perceptions of retrieval practice's effectiveness for retention are similar for oneself and others (McCabe, 2011), little is known about how its perceived impact on emotional well-being differs when considering others versus oneself. This adjustment provided critical insights into how learners regulate retrieval practice based on learning and well-being goals for themselves compared to others. I predicted a similar pattern of results, with a learning goal increasing the use of retrieval practice, while a focus on emotional well-being would discourage its use.

Experiment 4 aimed to explore the dynamics of students' motivations and strategy choices in the context of a semester-long course. Conducted in an undergraduate college classroom, this experiment examined how students' goals and strategy choices were influenced by the value of an upcoming exam and their awareness of the benefits of retrieval practice.

Research suggests that motivations are dynamic and subject to change over time to align with situational demands (Kim et al., 2023; Thomas et al., 2022). External factors, such as perceived task importance and value, are pivotal in shaping motivational priorities and subsequent learning decisions (Eccles et al., 1983; Wigfield & Eccles, 2000; Wigfield et al., 2009). In this experiment, participants took several unit exams throughout the semester. The weight of the exams was manipulated so that some exams accounted for 25% of the final grade, while others accounted for only 5%. I predicted that participants would be more motivated to learn and master the course material when expecting 25% weighted exams compared to 5% weighted exams (Liu et al., 2024).

In Liu et al. (2024), participants selected between retrieval practice and restudying to review exam topics. The long spacing interval between initial exposure and review may have led students to doubt their ability to retrieve information successfully, prompting them to prioritize restudying over self-testing (Son, 2005; Toppino et al., 2018), especially for high-stakes topics. However, in Experiment 4 of this study, participants selected a strategy for an interval practice activity that took place immediately after the corresponding lecture. Unlike Liu et al. (2024), I predicted that a larger proportion of students would choose retrieval practice over study practice when expecting 25% weighted exams, due to its learning benefits and higher expectations of retrieval success.

Similar to Experiment 2, I also isolated the effects of goals on strategy use by controlling the information provided to participants. In this experiment, this factor was manipulated within participants. Specifically, in the middle of the semester, participants received a series of memory lectures that disclosed the benefits of retrieval practice for long-term retention, supported by scientific evidence. Research suggests that such explicit instruction enhances recognition of retrieval practice's benefits (Ariel & Karpicke, 2018; McCabe, 2011). I predicted that more participants would use retrieval practice when highly motivated to learn, especially when they understood its benefits for learning.

Experiment 1A

Experiment 1 examined how goal orientations influenced the choice of study strategies by comparing the effects of prioritizing learning, prioritizing emotional well-being, or not having a stated priority on SRL. Previous research by Kornell and Son (2009) demonstrated that when learners were given the choice between practice testing and restudying for a vocabulary learning task, they tended to allocate more time to practice testing, using it as a diagnostic tool to identify areas for improvement. I anticipated replicating these findings in the current study, showing that participants in both the learning goal condition and the no goal condition would prioritize learning optimization by allocating more time to retrieval practice. In contrast, those in the well-being goal condition may exhibit hesitancy toward retrieval practice due to concerns about potential emotional costs, thus allocating less time to this strategy (Rea et al., 2022).

Further, Experiment 1A examined how goal orientations influenced the regulation of retrieval practice under varying learning circumstances by altering time limits for study and retrieval practice across two study sessions. In one condition, participants had more time for study practice compared to retrieval practice, while in the other condition they had more time for

retrieval practice. This manipulation was expected to widen the gap in participants' perceptions of anxiety and effectiveness between the two strategies. If the shorter time allotted for retrieval practice led to higher anxiety and a lower perceived benefit compared to study practice (Onwuegbuzie & Seaman, 2010), this contrast would be magnified under conditions of uneven time limits. These differences offer insights into goal-driven SRL by illustrating how goals shape learners' approaches to balancing emotional well-being and strategy effectiveness.

Method

Design

Experiment 1A employed a 2 × 2 × 3 mixed design. Both type of practice (retrieval practice vs. study practice) and time limit for each practice trial (5 seconds vs. 7 seconds) were manipulated within participants, while type of goal (no goal, learning goal, well-being goal) was manipulated between participants.

Participants

We conducted an a priori power analysis using G*Power (Faul et al., 2007) to determine the required sample size for our hierarchical regression models. Model 1 included control variables: retrieval practice performance, perceptions of effectiveness, and perceptions of anxiety and frustration. Model 2 introduced the main effect of goal orientation, with goal conditions dummy-coded into two predictors representing the three-level categorical variable (no goal, learning goal, and well-being goal). Model 3 added the main effect of practice trial time limit. This resulted in three tested predictors and a total of six predictors. The alpha level was set at .05, with power = .80. Given the limited empirical evidence on the effects of goal orientation on study time allocation, we assumed a small-to-medium effect size ($f^2 = 0.05$) for the power analysis, which suggested a required sample size of 223 participants. To account for potential

attrition, we recruited 302 participants via Prolific, an online participant recruitment platform. We recruited participants aged 18 to 30 years who were fluent in English. They received compensation of US$8 per hour.

Following predefined exclusion criteria, we removed one participant over age 30, two participants with moderate or extensive knowledge of Swahili, and two participants who did not follow instructions and frequently took notes on word pairs. An additional 11 participants were excluded for scoring below 50% accuracy across three rounds of study practice, indicating low attention. We also removed 26 participants who failed to correctly identify their assigned goal and 10 participants whose reported time allocations to retrieval and study practice did not sum to 100% (e.g., assigning 60% to retrieval practice and 70% to study practice). Finally, two participants were excluded for misunderstanding the study strategies, confusing the visual search task with retrieval practice.

After these exclusions, the final sample size was 248 participants (M_{age} = 24.32, SD_{age} = 3.52). Among them, 60.08% were female, 47.98% identified as White, 21.37% as Black, and 8.06% as Latino/Hispanic. Regarding education level, 19.35% were high school graduates, 32.66% held a bachelor's degree, and 11.29% had a master's degree or higher. There were 81 participants in the no goal condition, 82 in the learning goal condition, and 85 in the well-being goal condition. A sensitivity analysis using G*Power indicated that this sample size could detect a minimum effect size of $f^2 = 0.045$[1].

[1] The effect size of goal orientation observed in Experiment 1A was $f^2 = 0.039$, which would require a sample size of $N = 284$ to detect, indicating that the study was slightly underpowered. We addressed this in the subsequent experiments.

Materials

The learning materials consist of 20 Swahili-English word pairs (e.g., mashua – boat, malkia – queen), adapted from Karpicke and Roediger (2008). Ten pairs were allocated to the study practice session, and the remaining ten to the retrieval practice session. The wordlist assignment was counterbalanced across participants.

Procedure

Figure 1 illustrates the overall procedure of Experiments 1A. After expressing interest in participating, participants were directed to the Qualtrics platform where the study was conducted. At the beginning of the experiment, participants were briefed that they would undergo two study sessions, each involving the learning of ten word pairs. Each session consisted of an initial presentation of word pairs followed by three rounds of practice (either study practice or retrieval practice) to learn a set of made-up words and their corresponding English translations. To prevent external references, participants were not informed that the word pairs were Swahili–English.

Following the instructions, the first learning session began. Participants were presented with ten word pairs sequentially, each displayed for 5 seconds with a 0.5-second interval between pairs. They were instructed to read the pairs and attempt to learn them without taking notes. Afterward, participants completed three rounds of either retrieval practice with feedback or study practice. The order of the study strategies was counterbalanced.

For retrieval practice, participants were prompted to provide the English translation of a Swahili word within either 5 or 7 seconds (e.g., *mashua – ?*), followed by 2 seconds of feedback indicating the correctness of their answers, along with the correct answer (e.g., “Incorrect! The correct answer is: mashua – boat”). Feedback was color-coded for differentiation: “Correct”

appeared in blue, while “Incorrect!” appeared in red. A 0.5-second interval was incorporated before participants advanced to the next retrieval practice trial.

During study practice, participants were presented with Swahili-English word pairs sequentially and asked to enter the English translation of the Swahili cue within either 7 or 5 seconds (e.g., *malkia – queen; malkia – ?*). This maintains consistency in the activity format between study practice and retrieval practice, without requiring participants to exert additional effort to retrieve the English translation during study practice. Moreover, given the online platform, this activity ensures participant engagement in study practice. A 0.5-second interval was inserted between the study practice trials.

If the time limit for retrieval practice was set to 5 seconds, the time limit for study practice was set to 7 seconds, and vice versa. However, all participants were told that the time limit was consistently 5 seconds for both retrieval and study practice. This was done to ensure that the primary difference between the two study strategies – retrieving information from memory versus simply reviewing the material – was emphasized, without influencing participants' perceptions of the time constraints associated with each strategy. By keeping the instructions on the time limit consistent across both strategies, we aimed to focus participants' attention on the nature of the practice itself, rather than on any perceived difficulty arising from varying time limits.

Participants completed brief, 1-minute filler tasks at three points during the experiment: one after the initial learning phase and two following the first and second practice rounds (Figure 1). During each filler task, participants engaged in a visual search task. They were shown a photo containing hidden items, with a list of these items displayed at the bottom of the photo. Participants had one minute to study the photo, identify the hidden items, and report the number

of items they successfully located. The photos used for this task were sourced from Printalee (Printablee, n.d.).

After participants finished learning both sets of word pairs, they encountered one of three goal scenarios. They were asked to allocate practice time between retrieval practice and study practice strategies to help the hypothetical student achieve their goal. Participants were instructed to distribute the practice time as a percentage, ensuring the total added up to 100%. Under the learning goal scenario, participants received the following instructions:

> *A fellow student is studying new word pairs. They know that certain study strategies are effective for learning. Their primary goal is to learn as many word pairs as possible. After reviewing the word pairs once, they want to practice to stabilize learning. They are considering how frequently to incorporate retrieval practice and study practice into their practice routine. If you were in their position, how would you distribute the practice time between these two strategies?*

Under the well-being goal scenario, the following instructions were provided:

> *A fellow student is studying new word pairs. They know that certain study strategies may cause frustration and anxiety. Their primary goal is to safeguard their emotional well-being during their study sessions. After reviewing the word pairs once, they want to practice to stabilize learning. They are considering how frequently to incorporate retrieval practice and study practice into their practice routine. If you were in their position, how would you distribute the practice time between these two strategies?*

In the no goal condition, participants were not prompted to prioritize either goal, and there was no mention of the effects of study strategies on learning and emotion.

A fellow student is studying new word pairs. After reviewing the word pairs once, they want to practice to stabilize learning. They are considering how frequently to incorporate retrieval practice and study practice into their practice routine. If you were in their position, how would you distribute the practice time between these two strategies?

After completing the time distribution task, participants answered several manipulation check questions. They were asked to identify the primary goal of the hypothetical student in the scenario, selecting from three options: to safeguard emotional well-being, to learn as many word pairs as possible, or not specified. They were also asked to explain how they distributed practice time between retrieval practice and study practice for the hypothetical student, and why they believed that distribution would help the student achieve their goal. Additionally, participants were asked to report their perceptions of strategy effectiveness and anxiety by answering the following questions: "Which study strategy do you think is more effective for learning: retrieval practice or study practice?" and "What makes you believe that strategy is more effective for learning?" They were also asked, "Which study strategy do you think might induce anxiety and frustration while studying: retrieval practice or study practice?" and "What makes you believe that strategy might induce anxiety and frustration?"

Participants were asked to report whether they had ever written down or looked up the word pairs during the experiment, using a 5-point Likert scale (from 1 = never to 5 = always). Participants were assured that their responses would not affect their compensation. They were then informed that the word pairs they had learned were actually Swahili words and their English translations, and asked to report their prior knowledge of Swahili on a 5-point Likert scale (from 1 = none to 5 = extensive). Following this, participants completed a demographic questionnaire.

Upon completing the survey, participants were compensated at a rate of USD 8 per hour. The entire study lasted approximately 30 minutes.

Results

Performance During Retrieval Practice

Figure 2 shows participants' retrieval practice performance across rounds, broken down by goal conditions and practice trial time limit. A 2 (practice trial time limit: 5s vs. 7s) × 3 (type of goal: no goal, learning, well-being) × 3 (round: 1, 2, 3) mixed-design analysis of variance (ANOVA) was conducted on retrieval practice performance, with practice trial time limit and type of goal as between-participants factors and round as a within-subjects factor.

Results revealed a significant main effect of round, $F(2, 484) = 317.39$, $p < .001$, $\eta^2_g = .123$, indicating that retrieval accuracy improved across rounds. On average, accuracy was $M = 0.39$ ($SD = 0.30$) in round 1, $M = 0.56$ ($SD = 0.31$) in round 2, and $M = 0.66$ ($SD = 0.29$) in round 3. There was also a significant main effect of time limit, $F(1, 242) = 11.18$, $p = .001$, $\eta^2_g = .040$, indicating that retrieval accuracy was higher when participants had 7 seconds per trial ($M = 0.60$, $SD = 0.32$) compared to 5 seconds per trial ($M = 0.48$, $SD = 0.31$). The main effect of goal orientation was not significant, $F(2, 242) = 0.66$, $p = .516$, indicating that retrieval accuracy did not differ across goal conditions[2]. Additionally, none of the interaction effects, including time limit × goal orientation, time limit × round, or goal orientation × round, were significant ($ps > .05$).

[2] Retrieval practice performance was measured during the initial learning phase, before participants were exposed to the goal scenarios. Therefore, no differences in performance were expected across goal conditions.

Perceptions of Effectiveness, Anxiety and Frustration

Most participants perceived retrieval practice as more effective (63.31%) and more likely to induce anxiety and frustration (87.10%). Table 1 shows participants' perceptions of strategy effectiveness, anxiety, and frustration across goal conditions and practice trial time limits. *Chi-square* tests of independence showed that perceptions of effectiveness ($\chi^2(2) = 0.35$, $p = .839$) and perceptions of anxiety and frustration ($\chi^2(2) = 0.36$, $p = .834$) did not vary across goal conditions. Similarly, no significant differences were found in perceptions of effectiveness, ($\chi^2(1) = 1.54$, $p = .215$) or perceptions of anxiety and frustration ($\chi^2(1) = 1.93$, $p = .164$) across the two time limits.

Time Allocated to Retrieval Practice

Figure 3 shows the percentage of study time participants allocated to retrieval practice, as a function of goal condition and practice trial time limit. To examine predictors of time allocated to retrieval practice, we conducted a series of hierarchical linear regressions. Table 2 displays the summary of hierarchical regression analyses for predicting the time allocated to retrieval practice in Experiment 1A.

Model 1: Control Variables. In Model 1, I included retrieval practice performance during the final round of practice, along with participants' perceptions of effectiveness and of anxiety and frustration, as control variables. For the perception variables, selecting 'study practice' as more effective and more likely to induce anxiety and frustration served as the reference level. Retrieval practice performance on the final round was not a significant predictor of time allocated to retrieval practice, $B = 8.39$, $SE = 4.30$, $p = .053$. However, participants who perceived retrieval practice as more effective allocated significantly more time to it, $B = 21.67$,

SE = 2.60, $p < .001$. Perceptions of anxiety and frustration were not a significant predictor, *B* = -5.77, *SE* = 3.28, $p = .080$.

Model 2: Adding Type of Goal. In Model 2, I added type of goal as a predictor, using the no goal condition as the reference level. Participants assigned to the well-being goal condition allocated significantly less time on retrieval practice than those in the no goal condition, *B* = -8.35, *SE* = 2.60, $p = .002$. In contrast, participants in the learning goal condition did not differ significantly from those in the no goal condition, *B* = 0.91, *SE* = 2.63, $p = .729$. An *F*-test comparing Model 1 and Model 2 showed that incorporating type of goal into the model improved model fit, $F(2, 242) = 7.78$, $p < .001$, indicating that goal orientation contributed additional explanatory power beyond the control variables.

Model 3: Adding Time Limit. Next, I added time limit (5 seconds vs. 7 seconds per retrieval trial) as a predictor, with 5 seconds set as the reference level. The results showed that time limit was not a significant predictor of study time allocation, *B* = 0.52, *SE* = 2.19, $p = .812$. Further, adding time limit did not significantly improve the model fit, $F(1, 241) = 0.06$, $p = .812$.

Model 4: Adding Interaction Between Type of Goal and Perceived Effectiveness. In Model 4, we also explored the interaction between goal orientation and perceived effectiveness to examine whether the impact of goal orientation on time allocated to retrieval practice varied based on perceptions of strategy effectiveness. The interaction terms were not significant (*ps* > .05), indicating that the relationship between goal orientation and time allocation did not depend on perceptions of strategy effectiveness. Model comparison showed that adding the interaction did not improve model fit, $F(2, 239) = 0.83$, $p = .437$.

Discussion

Experiment 1A examined the effects of goal orientation (learning goal, well-being goal, or no goal) and practice trial time limit (5 seconds or 7 seconds per retrieval trial) on study time allocation to retrieval practice. The main finding was that the goal assigned to participants significantly influenced the amount of time they allocated to retrieval practice. Specifically, those in the well-being goal condition allocated significantly less time to retrieval practice compared to those in the no goal condition. No significant difference was found between the learning goal and no goal conditions, likely because most participants (86.42%) in the no goal condition believed the primary goal of the hypothetical student was to learn as many word pairs as possible, aligning with the learning goal. Additionally, the results confirmed our prediction that most participants recognized the trade-off between the learning benefits of retrieval practice and its negative emotional impacts, which helps explain why emotional well-being goals reduced engagement with retrieval practice compared to learning goals.

Time limit, on the other hand, did not significantly affect study time allocation, despite participants performing better when given 7 seconds per retrieval trial compared to 5 seconds. This lack of effect may be explained by the minimal difference in time limits, which did not significantly alter participants' perceptions of strategy effectiveness or levels of anxiety and frustration.

However, the varying time limits across goal conditions made it challenging to isolate the effect of goal orientation on study time allocation. Research indicates that participants are more likely to engage in retrieval practice when they experience greater success in retrieval (Son, 2005; Toppino et al., 2018). Although time limit was not a significant predictor of study time allocation, it did significantly impact retrieval practice performance, which may have, in turn,

influenced how some participants allocated study time between the two strategies—making it more difficult to interpret the effect of the goal conditions. To address this, Experiment 1B kept the time limit consistent across all goal conditions, allowing for a clearer examination of the effect of goal orientation on study time allocation without the influence of different time limits.

Experiment 1B

Experiment 1B aimed to replicate the findings of Experiment 1A in a design where time limit was consistent across goal conditions. I expected to observe a similar pattern of results, with participants allocating less time to retrieval practice when encouraged to prioritize well-being, and a similar amount of time to retrieval practice when encouraged to prioritize learning, compared to the no goal condition. This expectation was based on the finding from Experiment 1A that most participants in the no goal condition believed the hypothetical student's goal was to maximize learning.

Method

Design

Experiment 1B employed a 2 × 3 mixed-design, with type of practice (retrieval practice vs. study practice) manipulated within participants and type of goal (no goal, learning goal, well-being goal) manipulated between participants.

Participants

We conducted a power analysis using G*Power (Faul et al., 2007) to determine the required sample size for our incremental regression models. Model 1 was similar to the one used in Experiment 1A, while Model 2 incorporated dummy codes for type of goal, resulting in two additional predictors, for a total of five predictors. The analysis was performed with $\alpha = .05$, power = .80, and an effect size of $f^2 = 0.039$, based on Experiment 1A. G*Power estimated a

required sample size of 251 participants. However, to account for potential attrition, we recruited 353 participants, using the same data collection method and targeting the same population as in Experiment 1A.

Based on predefined exclusion criteria, we removed four participants over the age of 30. Three participant were excluded due to moderate or extensive knowledge of Swahili, and four participants were removed for not following instructions by frequently or always taking notes on the word pairs. An additional 14 participants were excluded for scoring below 50% accuracy across all three rounds of study practice, indicating low attention. We also excluded 20 participants who failed to correctly identify their assigned goal condition and 12 participants whose reported time allocations to retrieval and study practice did not sum to 100%.

After these exclusions, the final sample consisted of 296 participants (M_{age} = 24.95, SD_{age} = 3.44). Among them, 59.46% were female, 54.05% identified as White, 12.50% as Black, and 8.45% as Latino/Hispanic. Regarding education level, 15.88% were high school graduates, 35.81% held a bachelor's degree, and 14.19% had a master's degree or higher. Due to a technical issue with Qualtrics, there was an uneven distribution of participants across conditions: 84 in the no goal condition, 128 in the learning goal condition, and 84 in the well-being goal condition. A sensitivity analysis using G*Power showed that this sample size, with the smallest condition containing 84 participants, was sufficient to detect a minimum effect size of $f^2 = 0.039$.

Materials and Procedure

Experiment 1B used the same materials and overall procedure as Experiment 1A, with one key modification. During practice, participants were required to enter the English translations of the Swahili words within seven seconds per trial, regardless of the study strategy (retrieval or study practice).

Results

Performance During Retrieval Practice

Due to a programming error, one word pair was omitted from the data analysis in one of the two word lists. Figure 4 shows participants' retrieval practice across rounds and goal conditions. A 3 × 3 mixed-design ANOVA was conducted to evaluate the effects of practice round (within-subjects: round 1, round 2, round 3) and goal condition (between-subjects: no goal, learning, well-being) on retrieval accuracy. There was a significant main effect of practice round, $F(1.84, 586) = 326.74$, $p < 0.001$, $\eta^2_g = 0.121$, demonstrating that retrieval accuracy improved over three practice rounds. On average, retrieval accuracy was $M = 0.43$ ($SD = 0.31$) in round 1, $M = 0.61$ ($SD = 0.30$) in round 2, and $M = 0.70$ ($SD = 0.28$) in round 3. The main effect of goal condition was not significant, $F(2, 293) = 1.49$, $p = .228$, indicating that retrieval accuracy did not vary across goal conditions. Similarly, the interaction between goal condition and practice round was not significant, $F(3.70, 586) = 0.93$, $p = .444$, suggesting that retrieval performance improved across rounds in a comparable manner regardless of goal condition.

Perceptions of Effectiveness, Anxiety and Frustration

Most participants perceived retrieval practice as more effective (68.92%) and more likely to induce anxiety and frustration (87.16%). Table 3 shows participants' perceptions of strategy effectiveness, anxiety, and frustration across goal conditions. *Chi-square* tests of independence showed that perceptions of effectiveness, $\chi^2(2) = 3.60$, $p = .166$, and perceptions of anxiety and frustration, $\chi^2(2) = 0.87$, $p = .646$, did not vary across goal conditions.

Time Allocated to Retrieval Practice

Similar to Experiment 1A, a series of hierarchical linear regressions were performed to examine predictors of time allocated to retrieval practice. Table 4 displays the summary of

hierarchical regression analyses for predicting the time allocated to retrieval practice in Experiment 1B.

Model 1: Control Variables. In Model 1, retrieval practice performance (round 3), perceived effectiveness, and perceived anxiety and frustration were included as control variables. Results showed that retrieval practice performance did not significantly predict time allocated to retrieval practice, $B = 5.26$, $SE = 4.33$, $p = .225$. Participants who perceived retrieval practice as the more effective study strategy allocated significantly more time to it, $B = 22.86$, $SE = 2.59$, $p < .001$. However, perceptions of anxiety and frustration were not a significant predictor, $B = 3.65$, $SE = 3.36$, $p = .278$.

Model 2: Adding Type of Goal. In Model 2, dummy variables representing type of goal were added as predictors. Levene's test indicated no significant variance differences across goal conditions, $F(2, 293) = 0.140$, $p = .870$, confirming that regression assumptions are met despite unequal sample sizes. Results indicated that participants in the well-being goal condition allocated significantly less time to retrieval practice compared to those in the no goal condition, $B = -16.44$, $SE = 2.79$, $p < .001$. However, there was no significant difference between the learning goal and no goal conditions, $B = -2.24$, $SE = 2.54$, $p = .380$. An F-test comparing Model 1 and Model 2 showed that adding goal condition significantly improved model fit, $F(2, 290) = 21.23$, $p < .001$, suggesting that goal condition explains additional variance in time allocated to retrieval practice beyond the control variables. Figure 5 shows the percentage of study time participants allocated to retrieval practice, as a function of goal conditions.

Model 3: Adding Interaction Between Type of Goal and Perceived Effectiveness. In Model 3, we explored the interaction between type of goal and perceived strategy effectiveness

to assess whether the effect of goal orientation on time allocated to retrieval practice varied by perceived effectiveness.

Results showed that neither the interaction between perceived effectiveness and the well-being goal condition (B = -10.47, SE = 6.27, p = .096) nor the interaction between perceived effectiveness and the learning goal condition (B = 3.71, SE = 5.44, p = .495) was significant. However, a model comparison using an F-test revealed that adding the interaction significantly improved model fit, $F(2, 288) = 3.25$, p = .040, suggesting that the interaction between goal condition and perceived effectiveness explained additional variance in time allocated to retrieval practice.

For exploratory purposes, post-hoc comparisons were conducted to further examine time allocation across goal conditions based on participants' perceived strategy effectiveness. Among participants who viewed study practice as more effective, time allocation did not significantly differ across conditions ($ps > .05$). In contrast, among participants who believed that retrieval practice was more effective, significant differences in time allocation emerged across goal conditions. Participants in the well-being goal condition allocated less time to retrieval practice (M = 49.71, SD = 21.84) compared to those in the no goal condition (M = 68.40, SD = 16.03; B = -19.10, SE = 3.24, $p < .001$) and the learning goal condition (M = 67.68, SD = 14.34; B = -18.29, SE = 3.01, $p < .001$). No significant difference was found between the no goal and learning goal conditions (B = 0.81, SE = 3.05, p = 1.00). Figure 6 shows the percentage of study time participants allocated to retrieval practice, as a function of goal conditions and perceived effectiveness.

Discussion

Experiment 1B replicated and extended the findings from Experiment 1A. That is, under the constraints of identical time limits across goal conditions, participants who received the well-being goal allocated less time to retrieval practice than participants who did not receive any specific goal. No significant difference in study time allocation was observed between the learning goal and no goal conditions, likely because most participants in the no goal condition (80.95%) identified the assigned goal as a learning goal. This finding further supports the idea that many learners recognize the trade-off between the learning benefits of retrieval practice and the negative emotional experiences it can trigger. As a result, prioritizing emotional well-being reduces engagement with effortful learning strategies like retrieval practice, even when learners are aware of its benefits for learning.

Interestingly, the effect of goal orientation on study time allocation varied based on participants' perceptions of retrieval practice's effectiveness. Participants who viewed retrieval practice as more effective for learning allocated more time to it when prioritizing learning, compared to well-being. In contrast, those who preferred study practice did not show significant differences between goal conditions. This suggests that the impact of goal orientation on study time allocation is most noticeable when participants recognize the learning benefits of retrieval practice.

Experiment 1B provided a clearer understanding of how goals influence study time allocation by controlling for time limit differences. However, it still had limitations in fully isolating the effect of goal, as individual differences, such as perceptions of strategy effectiveness and retrieval practice performance, also influenced how participants allocated time between retrieval and study practice. Perceptions of effectiveness appeared to moderate the

relationship between goal and study time allocation. Additionally, although retrieval practice performance did not significantly impact study time allocation in Experiments 1A and 1B, research shows that learners are more likely to engage in retrieval practice when the likelihood of retrieval success is higher, such as with easier items or shorter spacing intervals (Son, 2005; Toppino et al., 2018).

Because individual differences like perceived effectiveness and retrieval practice performance were not controlled in Experiments 1A and 1B, our ability to draw clear conclusions about the effect of goal orientation was limited. To address these limitations, Experiment 2 was designed to control for these individual differences, allowing for a more accurate examination of the effect of goal orientation on study time allocation.

Experiment 2

In Experiment 2, I isolated the effects of retrieval practice performance and perceptions to better understand goal-driven SRL. To achieve this, participants did not directly experience either retrieval or study practice. Instead, they only received descriptions of the two study strategies. Next, some participants were provided with research information on the benefits of retrieval practice for learning and its potential negative impacts on emotional well-being. This condition was compared to a control group that received a passage of similar length, unrelated to the study. Participants then allocated study time between retrieval and study practice based on the assigned goal (learning goal or well-being goal). This design allowed me to examine the effect of goal orientation on study time allocation while controlling for individual differences, such as perceptions of strategy effectiveness and retrieval practice performance.

Method

Design

Experiment 2 employed a 2 × 2 between-participants design, manipulating two factors: type of goal (learning goal vs. well-being goal) and the provision of research information (research information: absent vs. provided). This design resulted in four conditions: learning goal, research absent; learning goal, research provided; well-being goal, research absent; and well-being goal, research provided.

Participants

We conducted a power analysis using G*Power (Faul et al., 2007) to estimate the necessary sample size for our incremental regression models. Model 1 included perceptions of effectiveness and perceptions of anxiety and frustration before exposure to the goal scenarios and research information as control variables. Model 2 introduced the main effect of goal condition, while Model 3 added the main effect of research condition. Finally, Model 4 added the interaction between goal and research condition, resulting in a total of three tested predictors and five predictors overall. The analysis was performed with $\alpha = .05$, power = .80, and an effect size of $f^2 = 0.039$, derived from Experiment 1A. We used the smallest observed effect size for goal condition in previous studies to ensure a conservative estimate. G*Power indicated a required sample size of 284 participants. However, to mitigate potential attrition, we recruited 343 participants. The data collection method and target population remained aligned with the prior experiments.

After data collection, 24 participants were excluded for failing to correctly identify their assigned goal condition. Additionally, 20 participants were excluded for failing to correctly identify retrieval practice as both effective and more likely to induce anxiety and frustration, as

stated in the provided research information[3]. Six participants were excluded for incorrectly answering more than 50% of the questions related to the passage information in the research-absent condition. This left a final sample of 293 participants (M_{age} = 25.00, SD_{age} = 3.40). Among them, 61.43% were female, 53.58% identified as White, 10.24% as Black, and 6.48% as Latino/Hispanic. In terms of education, 16.38% were high school graduates, 40.96% held a bachelor's degree, and 8.19% had a master's degree or higher.

Participants were distributed across conditions as follows: 68 in the learning-goal, research-absent condition; 73 in the learning-goal, research-provided condition; 80 in the well-being-goal, research-absent condition; and 72 in the well-being-goal, research-provided condition. A sensitivity analysis using G*Power indicated that this sample size could detect a minimum effect size of f^2 = 0.038.

Materials and Procedure

Figure 7 displays the overall procedure of Experiment 2. The experiment began with instructions asking participants to review information about common study strategies for learning vocabulary. They were instructed to read carefully, as they would be asked questions about it afterward. Next, participants were provided with descriptions of the two study strategies:

This study examines how students learn new vocabulary using two different strategies:

[3] These participants were removed because the questions assessed comprehension of the provided content—not participants' personal perceptions—so an incorrect response indicated inattention to the information.

- ***Retrieval practice:*** *This strategy involves repeatedly recalling word meanings from memory, such as by using flashcards or taking practice tests on each word's definition.*
- ***Study practice:*** *This strategy involves repeatedly reading and writing each word and its definition, such as by rereading the words or copying them into a notebook.*

Participants were then asked to describe the two strategies in their own words, followed by immediate feedback in the form of the original descriptions. This activity was intended to draw their attention to the descriptions once more; therefore, I did not assess their responses for accuracy or exclude participants based on how accurately they described the strategies. Following that, they chose which strategy they thought would cause more anxiety and frustration, and which one they thought would be more effective for learning new vocabulary, between retrieval and study practice. The questions were the same as in Experiments 1A and 1B, with one addition. Participants were asked to rate, on a 5-point scale, how much more likely they thought the strategy they chose (compared to the one they did not choose) would cause anxiety and frustration (1 = slightly more likely to 5 = extremely more likely). They were also asked to rate how much more effective they thought the strategy they chose was compared to the one they didn't choose (1 = slightly more effective to 5 = extremely more effective). These Likert scale items were added to improve the statistical sensitivity of the two predictors – perceived strategy effectiveness and perceived anxiety and frustration – by replacing the binary choice from Experiments 1A and 1B (retrieval vs. study practice) with a more detailed scale. This approach allows for a more precise relationship between these predictors and the time allocated to retrieval practice.

Next, participants either received research information on retrieval practice (the research-provided condition) or a control passage of similar length, unrelated to the study (the research-absent condition). The information provided was as follows:

Research-Provided Condition.

Research shows that retrieval practice, a study strategy where students actively recall information, has both advantages and drawbacks.

Pros: With the same amount of study time, students remember more when they use retrieval practice compared to study practice strategies like rereading or rewriting. Retrieval practice strengthens the connection between words and their meanings, supporting long-term learning.

Cons: However, retrieval practice is a more demanding approach than study practice. It requires greater effort and can lead to mistakes during the recall process. These mistakes may cause frustration and anxiety, making it a challenging strategy for students.

Research-Absent Condition.

Research reveals that over 7,000 languages are spoken worldwide, each with distinctive features such as vocabulary, grammar, and writing systems. For example, Chinese uses logographic characters, while Spanish and English use alphabets.

The number of speakers of these languages varies widely. English and Mandarin are spoken by millions, making them among the most widely used languages globally. In contrast, languages like Basque and Hawaiian are spoken by only a few hundred people.

This linguistic diversity reflects the richness of human culture, with each language providing unique perspectives shaped by the traditions and ideas of its speakers.

To ensure participants read and fully understood the information, those in the research-provided condition were asked, “According to scientific research, which study strategy is more effective for learning new vocabulary: retrieval practice or study practice?” and to explain why that strategy is more effective based on the provided information (open-ended question). They were also asked, “According to scientific research, which study strategy might induce anxiety or frustration when learning new vocabulary: retrieval practice or study practice?” and to explain why that strategy might cause anxiety or frustration. Afterward, participants rated, “How much more likely do you think [the selected strategy] is to cause anxiety and frustration compared to [the unselected strategy]?” and “How much more effective do you think [the selected strategy] is compared to [the unselected strategy]?” on a 5-point Likert scale.

In the research-absent condition, participants were asked six questions related to the information they received. The format of these questions was similar to those in the research-present condition: four multiple-choice questions (e.g., “According to the passage, approximately how many languages are spoken worldwide?” with options: 500, 2,000, 7,000, or 10,000) and two open-ended questions (e.g., “How do languages like Chinese and Spanish differ in their writing systems, as mentioned in the previous passage?”).

Participants were then presented with either the learning goal or well-being goal scenario (identical to those in Experiments 1A and 1B) and asked to allocate practice time between retrieval and study practice. After completing the time distribution task, participants answered manipulation check questions, including identifying the primary goal of the hypothetical student (safeguard emotional well-being, maximize learning, or unspecified) and explaining how and why they would allocate practice time between retrieval and study practice to help the student

achieve that goal. Finally, participants completed a demographic questionnaire. The entire experiment took approximately 10 minutes.

Results

Perceived Learning and Perceived Emotion

Most participants perceived retrieval practice as more effective (59.73%) and more likely to induce anxiety and frustration (69.97%), even without directly experiencing both study strategies or being exposed to research information on the effects of retrieval practice on learning and emotional experiences. Table 5 shows participants' perceptions of strategy effectiveness, anxiety, and frustration across goal and research provision conditions before participants received research information. To facilitate further statistical analyses, participants' ratings of effectiveness and anxiety/frustration were transformed. If participants selected retrieval practice as more effective or more likely to induce anxiety and frustration, their Likert ratings remained unchanged (1 to 5). If they selected study practice, their ratings were reversed (-1 to -5).

Prior Ratings Across Conditions. A 2 × 2 between-participants ANOVA was conducted to examine whether participants' perceived effectiveness and anxiety differed based on goal orientation (learning vs. well-being) and the provision of research information (research absent vs. research provided) prior to receiving research information. For effectiveness ratings, the main effect of goal orientation was not significant, $F(1, 289) = 1.14$, $p = .286$, indicating that participants' perceptions of effectiveness did not differ between the learning ($M = 0.92$, $SD = 3.28$) and well-being goal conditions ($M = 0.49$, $SD = 3.54$). Similarly, the main effect of research provision was not significant, $F(1, 289) = 0.22$, $p = .637$. The interaction between goal orientation and research provision was also nonsignificant, $F(1, 289) = 0.78$, $p = .379$.

For anxiety ratings, the analysis revealed a significant main effect of goal orientation, $F(1, 289) = 4.25$, $p = .040$, indicating an unexpected difference in perceptions of anxiety and frustration between the two goal conditions. Specifically, participants in the learning goal condition reported higher anxiety ratings toward retrieval practice ($M = 1.64$, $SD = 2.65$) compared to those in the well-being goal condition ($M = 0.93$, $SD = 3.16$). In contrast, the main effect of research provision was not significant, $F(1, 289) = 1.27$, $p = .261$, nor was the interaction between goal and research provision, $F(1, 289) = 1.06$, $p = .304$.

Although there was an unexpected difference in anxiety and frustration ratings, this should not pose a problem for hypothesis testing, as these factors were included as control variables in the regression models.

Changes After Research Exposure. To assess whether participants' perceptions of effectiveness and perceptions of anxiety and frustration changed after receiving research information, I conducted paired-samples *t*-tests for effectiveness and anxiety ratings in the research-provided condition. For effectiveness ratings, the analysis revealed a significant increase in perceived effectiveness after participants received research information, $t(144) = -9.35$, $p < .001$. Specifically, mean effectiveness ratings increased from 1.48 ($SD = 2.80$) before receiving the research information to 3.23 ($SD = 1.20$) after receiving it.

For anxiety ratings, the analysis also revealed a significant increase in perceived anxiety and frustration toward retrieval practice after receiving research information, $t(144) = -8.34$, $p < .001$. Specifically, mean ratings of anxiety and frustration increased from 1.48 ($SD = 2.80$) prior to receiving the research information to 3.23 ($SD = 1.20$) after receiving it.

Time Allocated to Retrieval Practice

To examine predictors of time allocated to retrieval practice, a series of hierarchical linear regressions were performed. Table 6 displays the summary of hierarchical regression analyses for predicting the time allocated to retrieval practice in Experiment 2.

Model 1: Control Variables. I conducted a linear regression analysis to examine whether participants' prior perceptions of effectiveness and prior perceptions of anxiety and frustration toward retrieval practice predicted the time they allocated to using the strategy. The results showed that perceiving retrieval practice as more effective before research exposure was associated with increased time allocated to the strategy ($B = 2.70$, $SE = 0.38$, $p < .001$). In contrast, perceptions of anxiety and frustration were not a significant predictor ($B = -0.24$, $SE = 0.44$, $p = .585$).

Model 2: Adding Type of Goal. Building on Model 1, I added type of goal as a predictor. Perceiving retrieval practice as more effective remained a significant predictor of retrieval practice time ($B = 2.33$, $SE = 0.32$, $p < .001$), while perceiving it as anxiety-inducing was now a significant negative predictor ($B = -0.85$, $SE = 0.37$, $p = .023$). Figure 8 shows the predicted percentage of time participants allocated to retrieval practice based on their perceptions of strategy effectiveness, anxiety and frustration.

Participants in the well-being goal condition allocated significantly less time to retrieval practice than those in the learning goal condition ($B = -24.15$, $SE = 2.11$, $p < .001$). An F-test comparing Model 1 to Model 2 showed that adding type of goal significantly improved model fit, $F(1, 289) = 130.59$, $p < .001$, indicating that goal orientation accounted for a substantial portion of the variance in retrieval practice time beyond the control variables.

Model 3: Adding the Effect of Research Provision. Next, the two research provision conditions were incorporated into the model. Results showed that research provision was not a significant predictor of time allocated to retrieval practice (B = -3.40, SE = 2.09, p = .105). Moreover, including the effect of research provision did not lead to a significant improvement in model fit, $F(2, 288) = 2.65$, p = .105.

Model 4: Adding Interaction Between Type of Goal and Research Provision. In Model 4, I added the interaction between type of goal and research provision to examine whether the effect of goal orientation on time allocated to retrieval practice depended on research provision. Results revealed a significant type of goal × research provision interaction (B = -17.31, SE = 4.07, $p < .001$), indicating that the effect of research provision on retrieval practice time depended on the assigned goal. A model comparison suggests that adding this interaction significantly improved model fit relative to models including only the main effects, $F(1, 287) = 18.10$, $p < .001$.

Pairwise comparisons further clarified this interaction. Participants allocated more time to retrieval practice when encouraged to prioritize learning compared to when encouraged to prioritize well-being, regardless of whether research information on retrieval practice was provided ($ps < .05$). Within the well-being goal condition, participants who received research information allocated significantly less time to retrieval practice compared to those who did not (B = 11.7, SE = 2.81, $p < .001$). In contrast, within the learning goal condition, research provision had no significant effect on retrieval practice time (B = -5.6, SE = 2.93, p = .225). These findings suggest that research provision further reduced retrieval practice time for those encouraged to prioritize well-being, while having no significant impact on those encouraged to prioritize

learning. Figure 9 shows the percentage of study time participants allocated to retrieval practice, as a function of goal conditions and research provisions.

Discussion

Experiment 2 aimed to isolate the effects of goal orientation (learning goal vs. well-being goal) on study time allocation to retrieval practice, while controlling for individual differences such as perceptions of strategy effectiveness and retrieval practice performance. The results of Experiment 2 provided valuable insights into how goal-driven SRL operates, with key findings aligned with those of Experiments 1A and 1B.

Consistent with the findings from Experiments 1A and 1B, participants who were encouraged to prioritize emotional well-being allocated significantly less time to retrieval practice than those who prioritized learning. This supports the idea that prioritizing emotional well-being leads to reduced engagement with effortful learning strategies like retrieval practice.

In addition to goal orientation, the results suggest that perceptions of strategy effectiveness and emotional impact also play a significant role in study time allocation. As in Experiments 1A and 1B, perceiving retrieval practice as more effective for learning was associated with allocating more time to this strategy. However, unlike in previous experiments, higher perceived anxiety and frustration related to retrieval practice were linked to reduced time allocation in Experiment 2. This may be due to the use of a more nuanced measure of emotional cost. Participants not only chose the strategy they believed was more likely to induce anxiety and frustration, but also rated the degree to which they believed it would do so on a 5-point Likert scale. This approach improved the predictive value of emotional cost perceptions. In contrast, Experiments 1A and 1B used a binary measure of emotional cost. Because most participants

viewed retrieval practice as more likely to induce anxiety and frustration, the binary measure may have lacked variability, limiting its ability to predict strategy choice.

Another notable finding in this experiment was the interaction between goal conditions and research provisions. While research information about the benefits and emotional costs of retrieval practice did not significantly impact time allocation in the learning goal condition, it had a significant negative effect in the well-being goal condition. Specifically, participants in the well-being goal condition who received research information allocated significantly less time to using it. This could be because participants in Experiment 2 did not directly experience retrieval practice, making its emotional costs less salient to them without the research information. Exposure to research on the emotional costs of retrieval practice may have highlighted the potential negative emotions, such as anxiety and frustration, which likely reduced engagement with this strategy. In contrast, for those focused on learning, research exposure did not appear to alter their behavior, possibly because the time allocated to retrieval practice was already high without the exposure to research information (M = 59.84%, SD = 19.44). As a result, providing research information on the benefits of retrieval practice alone was not sufficient to further increase study time allocation to retrieval practice (M = 65.14%, SD = 19.51).

These findings align with previous research on goal-driven SRL, which suggests that learners' goals and perceptions, associated with both cognitive and emotional factors, play a significant role in shaping their strategy use (Boekaerts, 2011; Thomas et al., 2022). This study highlights that even when participants are made aware of the learning benefits of retrieval practice, emotional factors related to perceived anxiety and frustration can significantly influence their strategy choices, especially when their primary goal is emotional well-being.

Experiment 3

Experiments 1A, 1B, and 2 consistently demonstrated the robust effects of learning and well-being goals on how participants allocate study time for a hypothetical student. Building on these findings, Experiment 3 extended the investigation by examining how goal condition influences study time allocation when participants are asked to distribute study time for themselves. This shift from a hypothetical scenario to personal study behaviors allowed us to explore the broader implications of goal condition on SRL, providing valuable insights into how goal orientation shapes actual learning behaviors.

Method

Design

Experiment 3 employed a 2 × 3 mixed-design, with type of practice (retrieval practice vs. study practice) manipulated within participants and type of goal (no goal, learning goal, well-being goal) manipulated between participants.

Participants

We conducted a power analysis using G*Power (Faul et al., 2007) to determine the required sample size for our incremental regression models. All the three models were identical to Experiment 1B. The alpha level was set at .05, with power = .80, and an effect size of f^2 = 0.039 computed from Experiment 1A as it provided the smallest effect size of goal condition observed in prior experiments. G*Power estimated a required sample size of 251 participants. However, to account for potential attrition, we recruited 292 participants. The data collection method and target population were consistent with previous experiments.

Based on predefined exclusion criteria, we removed two participants over the age of 30, nine participants with moderate or extensive knowledge of Swahili, and six participants who

failed to follow instructions and took notes on the word pairs. Additionally, 14 participants were excluded for scoring below 50% accuracy across three rounds of study practice, indicating low attention, and 29 participants were removed for failing to correctly identify their assigned goal condition.

After these exclusions, the final sample consisted of 284 participants (M_{age} = 25.31, SD_{age} = 3.36). Among them, 57.39% were female, 60.92% identified as White, 14.79% as Black, and 4.58% as Latino/Hispanic. Regarding education level, 17.96% were high school graduates, 36.62% held a bachelor's degree, and 15.49% had a master's degree or higher. There were 93 participants in the no goal condition, 95 in the learning goal condition, and 96 in the well-being goal condition. A sensitivity analysis using G*Power indicated that this sample size could detect a minimum effect size of $f^2 = 0.034$.

Materials and Procedure

Experiment 3 followed the same materials and overall procedure as Experiment 1B, with the key difference that participants were asked to allocate practice time between retrieval and study practice for themselves. The goal scenarios were revised as follows:

The learning goal scenario. *Image you are continuing to study new word pairs. Think about which study strategies might be effective for your own learning. Your primary goal is to learn as many word pairs as possible. After reviewing the word pairs once, you need to practice to stabilize learning. How would you distribute your practice time between retrieval practice and study practice?*

The well-being goal scenario. *Image you are continuing to study new word pairs. Think about which study strategies might cause you frustration and anxiety. Your primary goal is to safeguard your emotional well-being during the study sessions. After reviewing the word pairs*

once, you need to practice to stabilize learning. How would you distribute your practice time between retrieval practice and study practice?

The no goal scenario. *Image you are continuing to study new word pairs. After reviewing the word pairs once, you need to practice to stabilize learning. How would you distribute your practice time between retrieval practice and study practice?*

Results

Performance During Retrieval Practice

Figure 10 displays retrieval practice performance across rounds and goal conditions. A 3 × 3 mixed-design ANOVA was conducted to examine the effects of practice round (within-subjects: round 1, round 2, round 3) and type of goal (between-subjects: no goal, learning, well-being) on retrieval practice performance. There was a significant main effect of practice round, $F(1.84, 586) = 326.74$, $p < .001$, $\eta^2_g = .121$, indicating that retrieval practice performance improved across rounds. On average, retrieval accuracy was $M = 0.43$ ($SD = 0.31$) in round 1, $M = 0.61$ ($SD = 0.30$) in round 2, and $M = 0.70$ ($SD = 0.28$) in round 3. However, the main effect of goal orientation was not significant, $F(2, 293) = 1.49$, $p = .228$, suggesting that retrieval practice performance did not differ across goal conditions. Additionally, the interaction between goal orientation and practice round was not significant, $F(3.70, 586) = 0.93$, $p = .444$, indicating that improvements in retrieval practice performance across rounds were similar across all goal conditions.

Perceptions of Effectiveness, Anxiety and Frustration

Most participants perceived retrieval practice as more effective (63.73%) and more likely to induce anxiety and frustration (82.39%). Table 7 shows participants' perceptions of strategy effectiveness, anxiety, and frustration across goal conditions. *Chi-square* tests of independence

showed that perceptions of effectiveness, $\chi^2(2) = 0.46$, $p = .793$, and perceptions of anxiety and frustration, $\chi^2(2) = 1.31$, $p = .521$, did not vary across goal conditions.

Time Allocated to Retrieval Practice

A series of hierarchical linear regressions, similar to those in Experiment 1B, were conducted to examine the predictors of time allocated to retrieval practice. Table 8 presents a summary of the hierarchical regression analyses predicting time allocated to retrieval practice in Experiment 3.

Model 1: Control Variables. In Model 1, I included retrieval practice performance (round 3), perceptions of effectiveness, and perceptions of anxiety and frustration as control variables. Results showed that participants who performed better on the last round of retrieval practice allocated more time to retrieval practice, $B = 17.67$, $SE = 4.48$, $p < .001$. Additionally, participants who believed retrieval practice was the more effective study strategy also allocated more time to retrieval practice, $B = 18.27$, $SE = 2.69$, $p < .001$. However, perceptions of anxiety and frustration did not significantly predict time allocated to retrieval practice, $B = -4.09$, $SE = 3.20$, $p = .202$.

Model 2: Adding Type of Goal. In Model 2, dummy variables for type of goal were added as predictors. Results showed that participants in the well-being goal condition allocated significantly less study time to retrieval practice compared to those in the no goal condition, $B = -10.16$, $SE = 2.90$, $p < .001$. However, there was no significant difference between the learning goal condition and the no goal condition, $B = -0.45$, $SE = 2.91$, $p = .877$. An F-test showed that adding type of goal to the model significantly improved model fit, $F(2, 278) = 7.94$, $p < .001$, indicating that goal orientation explains additional variance in retrieval practice time beyond the

control variables. Figure 11 shows the percentage of study time participants allocated to retrieval practice, as a function of goal conditions.

Model 3: Adding Interaction Between Type of Goal and Perceived Effectiveness. In Model 3, I included the interaction between goal orientation and perceived effectiveness to explore whether the effect of goal orientation on time allocated to retrieval practice varied by perceived strategy effectiveness. The results indicated that the interaction between perceived effectiveness and the well-being goal condition was not significant (B = -10.14, SE = 6.03, p = .094), nor was the interaction between perceived effectiveness and the learning goal condition (B = 5.13, SE = 6.02, p = .395). However, an F-test comparing model fit indicated that adding the interaction terms in Model 3 significantly improved the model compared to Model 2, $F(2, 276) = 3.44$, p = .033. This suggests that including the interaction between goal orientation and perceived strategy effectiveness enhanced the model's ability to explain the data.

Post-hoc analyses were conducted to explore differences in time allocated to retrieval practice across goal conditions depending on participants' perceptions of strategy effectiveness. Among participants who perceived study practice as more effective, time allocation did not significantly differ across goal conditions (ps > .05). In contrast, for participants who believed retrieval practice was more effective, significant differences emerged. Participants in the well-being-goal condition allocated significantly less time to retrieval practice (M = 50.94, SD = 19.58) compared to both the no goal (M = 65.49, SD = 17.93; B = 15.35, SE = 3.60, p = .002) and learning goal conditions (M = 65.62, SD = 17.60; B = 13.49, SE = 3.56, p = .002). No significant difference was found between the learning goal and no goal conditions (B = 1.61, SE = 3.63, p = .998). Figure 12 shows the percentage of study time participants allocated to retrieval practice, as a function of goal conditions and perceived effectiveness.

Discussion

The primary goal of Experiment 3 was to examine how goal orientation (learning goal, well-being goal, or no goal) influenced the allocation of study time to retrieval practice. Unlike Experiments 1A, 1B, and 2, where participants allocated time for a hypothetical student, participants in Experiment 3 were asked to allocate study time for their own learning, providing a more direct measure of how goal orientation influenced study behavior.

Consistent with previous findings, most participants perceived retrieval practice as more effective, but also as anxiety-inducing. This highlights the trade-off learners face when deciding whether to engage with effortful strategies like retrieval practice, which can be effective for learning but also emotionally taxing. Participants allocated significantly less time to retrieval practice when encouraged to safeguard emotional well-being, compared to those in the no goal condition. This mirrors findings from prior experiments, where emotional well-being goals led to reduced engagement with effortful strategies. There was no significant difference in retrieval practice time between participants in the no goal condition and those in the learning goal condition. This was likely because, similar to prior experiments, a large majority (84.95%) of participants in the no goal condition believed that the assigned goal was to maximize learning.

For the first time across Experiments 1-3, retrieval practice performance was found to significantly predict the time allocated to it. Participants were more likely to allocate additional time to retrieval practice as they retrieved more items during previous practice. This finding is particularly notable because, in Experiment 3, participants allocated time for their own study, allowing them to take their own performance into account. In contrast, in the previous experiments, participants allocated time on behalf of a hypothetical student, which likely resulted in more objective, less personal decision-making. This highlights the importance of retrieval

performance in shaping strategy choices when learners are directly involved in regulating their own study. These results align with prior research, which suggests that learners tend to engage more with retrieval practice when they believe they can successfully retrieve the studied information (Son, 2005; Toppino et al., 2018).

As in previous experiments, perceived effectiveness of retrieval practice was a significant predictor of time allocated to retrieval practice. Participants who perceived retrieval practice as more effective for learning allocated more time to it, reinforcing the role of metacognitive beliefs in strategy regulation.

More importantly, the impact of goal orientation appeared to vary depending on perceived effectiveness, mirroring findings from Experiment 1B. The influence of goal conditions on study time allocation to retrieval practice was only significant when participants perceived retrieval practice as more effective for learning. This suggests that when learners recognize the benefits of retrieval practice, their goal orientation (learning vs. well-being) becomes a more influential factor in guiding their study decisions. However, when participants viewed study practice as more effective, no significant differences were observed in time allocation across the goal conditions, underscoring the importance of perceptions of strategy effectiveness in shaping study behavior.

What distinguishes Experiment 3 from Experiments 1A, 1B, and 2 is that participants allocated time for their own study rather than for a hypothetical student. This shift allowed for a more direct examination of how individual goal orientation influenced personal study behavior, making these findings particularly relevant to real-world learning contexts. However, despite allocating time to retrieval and study practice for further learning, participants did not actually spend extra time on these strategies. Research suggests that the intention to use a strategy may

differ from its actual use, especially when there are perceived costs associated with its use (Wang et al., 2023). While students may intend to use retrieval practice to enhance learning, they may hesitate to engage in it when emotional costs, such as anxiety or frustration, are perceived to be high.

In Experiment 4, I brought my research question regarding the effect of goal orientation on study time allocation into the actual classroom environment, where students' goals, learning experiences, and perceptions were likely to interact and influence their use of study strategies. This experiment aimed to explore how learners navigated the trade-off between the learning benefits of retrieval practice and its emotional costs, and how they regulate their strategy choices in a real-world learning context.

Experiment 4

Experiment 4 examined how students' study strategy choices varied based on their learning and well-being goals in an in-person classroom setting. Prior research suggests that students' motivation to learn can influence how they engage with course material (Flake, 2015; Rea et al., 2022; Weissgerber et al., 2018). For instance, when students are more focused on learning, they may place greater value on retrieval practice. In contrast, when they prioritize reducing stress and avoiding negative emotional experiences, they may be less inclined to engage in effortful strategies such as retrieval practice.

Experiment 4 was conducted throughout the Fall 2024 semester in the Cognitive Psychology course. The course structure included bi-weekly unit exams. Exam weights were systematically manipulated so that some exams were more valuable than others. At the start of each class, participants indicated their learning and emotional well-being goals, followed by a

choice between two types of interval practice: interval testing (a form of retrieval practice) and interval study (a form of study practice).

Building on the work of Liu et al. (2024), I hypothesized that participants would be more motivated to learn when facing exams with higher weights compared to those with lower weights. Research indicates that heightened motivation to learn specific material leads to greater effort in mastering it (Weissgerber et al., 2016). Thus, I predicted that as exam weight increased, both learning goals and the adoption of retrieval practice would also increase. This prediction was different from the findings of Liu et al. (2024). In their study, participants selected a study strategy during a review session that occurred some time after the initial exposure. In contrast, in Experiment 4, participants selected a strategy type for a practice activity that took place immediately after the corresponding lecture, thereby reducing the likelihood of forgetting the content compared to the delayed strategy selection in Liu et al. (2024). Participants should be more likely to choose retrieval practice when they expect greater retrieval success (Son, 2005; Toppino et al., 2018). While I did not have a strong hypothesis regarding the impact of exam weight on emotional well-being goals, I speculated that participants might prioritize their well-being less when facing exams with higher stakes.

Additionally, as demonstrated in prior experiments, participants' perceptions of strategy effectiveness play a crucial role in shaping their strategy choices. To address this, similar to Experiment 2, information on the benefits of retrieval practice was provided through a series of memory lectures, which began midway through the semester. I predicted that more participants would choose interval testing after receiving the memory lectures (McCabe, 2011). Moreover, I anticipated an interaction between learning goals and memory lecture exposure. Those who were not initially aware of the benefits of retrieval practice for long-term memory might be less likely

to choose retrieval practice, even if highly motivated to learn. However, after receiving the memory lecture, participants with higher motivation to learn may be more likely to select interval testing over interval study.

Method

Design

The present study utilized a 2 (exam weight: 5%, 25%) × 2 (memory lecture: pre-exposure, post-exposure) × 17 (survey time points) within-subjects design. Each participant completed multiple unit exams throughout the semester, with each exam weighted either 25% or 5% of the final grade. After completing Survey 10, participants received memory lectures that explicitly discussed the benefits of retrieval practice in enhancing long-term retention of the studied material.

Participants completed 17 surveys throughout the semester, with each survey administered at the beginning of a class. In each survey, participants were asked to report their learning and well-being goals for that session. They also selected between two types of interval practice (interval study or interval testing) that they would engage in during the class. This design allowed us to track how their choices evolved over time and in response to their learning and well-being goals, the exam weights, and the memory lectures.

Participants

The participants were 35 undergraduate students enrolled in the Cognitive Psychology course at Tufts University during the Fall 2024 semester. Of these, 22 were female (62.86%). In terms of class year, five were freshmen (14.29%), 16 were sophomores (45.71%), nine were juniors (25.71%), and five were seniors (14.29%). Three students were granted special accommodations upon request. They completed the unit exams in a quiet, distraction-free lab

environment instead of the regular classroom and were allowed 1.5 times the standard time to complete each exam.

Materials

Pre-class survey. Participants completed this survey before attending the first class. In this survey, students reported their study strategies, self-testing habits, and levels of test anxiety. To assess study strategies, participants responded to the open-ended prompt, "How do you typically study on your own? Please describe all of the study strategies you regularly use." Self-testing frequency was measured with the item, "How often do you test yourself when you study on your own?" Participants rated their self-testing habits on a 4-point Likert scale (1 = almost never, 4 = almost always). Test anxiety was measured using the Test Anxiety Inventory (TAI) (Spielberger, 1980), which consists of 20 items (e.g., "During tests, I feel very tense"). Participants responded to each item on a 4-point Likert scale (1 = almost never, 4 = almost always), with higher scores indicating higher levels of test anxiety. The pre-class survey is provided in Appendix A.

Learning and well-being goals. These goals were measured on 5-point Likert scales, with questions including: "To what extent do you feel motivated to learn new materials?" and "To what extent do you want to relax and enjoy the class?". Higher points denote stronger endorsement (i.e., from 1 = not at all to 5 = extremely).

Interval practice. Participants engaged in interval practice during every class, choosing between two formats: interval study and interval testing. These choices were made at the beginning of the class, before exposure to the lecture. Regardless of the format, the interval practice consisted of 6 to 8 fill-in-the-blank questions, covering both the material from that day's lecture and the associated chapter from Chun and Most's (2021) *Cognition* textbook. The

questions were administered via Qualtrics and presented one by one to the students. Participants had 30 seconds to enter the words they believed best completed the blanks.

The key difference between the two formats was that in interval study, the correct answer was displayed on the same page as the question (e.g., “The mind can store encoded information about the environment as ____.” “Enter the answer: mental representations”). In interval testing, only the question was provided, and participants had to recall the answer. In both formats, the correct answer was shown afterward, embedded in the sentence with the answer in bold (e.g., “The mind can store encoded information about the environment as **mental representations**”).

Additionally, participants were asked to self-evaluate the accuracy of their responses by selecting one of three options to the question: “Did you get this question right?” They could select "Yes (1 point)," "Partially (0.5 point)," or "No (0 point)." After completing all the questions, participants received their total score based on their self-evaluations. This self-evaluation was intended for students to track their learning progress. Interval practice was graded based on completion only, not accuracy.

Mid-semester survey. Participants completed this survey after Exam 3. In this survey, they reported the number of hours spent on the class outside of regular class meetings by answering the question, "How many hours per week, outside of regularly scheduled class meetings, do you spend on this class?" Participants selected one of the following options: "Less than 1 hour," "1 – 2 hours," "3 – 4 hours," or "5 or more hours." They also reported the study strategies they used outside of class by responding to the question, "What study strategies do you use outside of class meetings to learn the course material?"

Additionally, participants reported which practice format (interval study or interval testing) they chose more often and the reasons for their choice. Specifically, they answered the

questions, "During this course, you have had the option to participate in either interval study or interval testing. Which strategy did you choose more often?" and "Why did you choose [the selected strategy] more often?"

Participants also assessed their perceptions of the effectiveness of the strategies, as well as their perceptions of anxiety and frustration. These questions were similar to those in Experiment 2, where participants indicated which strategy they believed caused more anxiety and frustration, and which they thought was more effective for learning between interval study and interval testing. Finally, they rated how much more effective and how much more likely they thought the chosen strategy would cause anxiety and frustration on 5-point Likert scales. The mid-semester survey is provided in Appendix B.

Course syllabus. The course syllabus provides students with a comprehensive outline, including a description of the interval practice, the schedule for the content covered in each class, and the schedule for unit exams with their associated weights (5% or 25% of the final grade). It also outlines the grading criteria for final grades. The course syllabus is provided in Appendix C. Students were encouraged to review the necessary material before each class meeting. In each in-class session, five minutes were allocated for the interval practice activity. Additionally, each class participation contributed 1% toward the final grade to encourage consistent engagement throughout the semester.

Procedure

The experiment was integrated into the Cognitive Psychology course at Tufts University. The course provided students with a comprehensive overview of both classic and contemporary topics in human cognition, covering a wide range of subjects from behavioral,

neuropsychological, and neuroscientific perspectives. Topics included the history of psychology, sensation, perception, attention, knowledge, memory, decision-making, and reasoning.

Table 9 displayed an overview of the course schedule. Taking place during the Fall 2024 semester, from September 2 to December 18, the course consisted of 13 weeks of class meetings held on Mondays and Wednesdays from 4:30 to 5:45 pm, with exceptions for university holidays on October 14 and November 11. All classes were conducted in person, except for the final three lectures, which were held remotely on November 25, December 2, and December 4, due to the instructor's schedule.

On the first day of class, all students received the course syllabus and informed consent. After receiving an overview of the course structure, students were asked to provide informed consent for their responses, including exam performance, strategy choice, and goals, to be used as data after the semester ended. To prevent coercion, informed consent forms were collected by a student volunteer at the end of the first class and sealed in an envelope delivered to the Psychology Department administrator. The envelope was opened by the course instructor and teaching assistant after final grades were posted, ensuring their unawareness of consent status until the course concludes.

The first and second classes were designed to familiarize participants with the class structure. Participants experienced the short lecture format and had the opportunity to engage in both interval study and interval testing. In the first class on September 4, they participated in interval study. In the second class on September 9, they experienced both interval testing and interval study. Beginning with the third class on September 11, participants completed a survey where they indicated their learning and well-being goals and selected their preferred interval practice format. Due to a technical issue, participants were unable to complete this survey during

the class meeting on September 16. However, they completed the survey in all subsequent classes, except on exam days.

In a standard lecture class, participants attended in person, completed the survey on Qualtrics, an online survey platform, listened to a 30-minute lecture, then engaged in a 5-minute interval practice activity, either through interval testing (retrieval practice) or interval study (study practice). They then listened to another 30-minute lecture and participated in another 5-minute interval practice activity, either interval testing or interval study. Figure 13 illustrates the class structure of a typical lecture session in Experiment 4. Interval practice activities, regardless of the format, were administered via Qualtrics. All study materials were made digitally accessible online the following day and remained available for the remainder of the semester. Due to limited class time, there were occasional instances when participants were unable to complete the second interval practice activity. However, these occurrences were random and unpredictable, and should not have impacted participants' reported goals or their choice of practice format at the beginning of the class.

There were three remote, asynchronous class meetings throughout the semester. They occurred on November 20, November 25, and December 2. For these sessions, the course instructor pre-recorded the lectures, and students were able to access the recordings through a Qualtrics survey. The same survey also included the standard questionnaire on learning and well-being goals, along with the choice of practice type. Participants were asked to complete the survey, watch the lectures, and engage in the interval practice activities on the same day as the scheduled class meeting.

There were bi-weekly unit exams, covering content from the most recent two weeks of lectures. The exams included multiple-choice questions (2 points each), fill-in-the-blank

questions (2 points each), short-answer questions (3 points each), and one essay question (10 points). While the topics overlapped with those in the interval practice activities, the exam questions were not repeated verbatim. Before completing each survey at the beginning of class, participants were reminded of the upcoming exam date and its weight (5% or 25%) in the final grade. Over the course of the semester, five unit exams were given. The first, third, and fifth exams each accounted for 5% of the final grade, while the second and fourth exams were each worth 25%. Participants had the option to choose between their Exam 1 and Exam 5 scores for the final grade calculation. If they selected Exam 1, they were not required to take Exam 5. On exam days, participants used the entire class time to complete the exam.

After each exam (with the exception of Exam 5) participants completed a take-home class activity related to an upcoming course topic. For example, they were introduced to the concept of childhood amnesia and asked to reflect on it by answering questions about their own childhood memories. Each activity was worth 5% of the final grade. With four activities in total, following Exams 1 through 4, these assignments accounted for 20% of the final grade. However, because these activities were part of the course rather than the experiment, they will not be discussed further.

For the class before each unit exam, the instructor provided approximately 30 minutes of lecture, followed by interval practice. Afterward, a review session was conducted, covering all material from the past two weeks that would appear on the upcoming exam (Figure 13). During the review session, participants were given several questions to discuss in small groups of two to three students. Each group was responsible for one question and had 10 minutes to discuss and prepare their answer, which was then presented to the whole class. Due to time constraints, the review session was not held for Exam 1 or Exam 5. For Exam 4, the review questions were

posted online one week in advance to give participants more time to prepare. After the review sessions, a study guide containing answers to all the review questions was made available online to help facilitate participants' preparation for the exam. Figure 13 illustrates the class structure for both a standard lecture class and a review class. To encourage attendance, participants received one credit (1%) for each class attended, excluding exam days. With 20 non-exam class meetings, attendance accounted for 20% of the final grade.

Results

Table 10 shows the descriptive statistics for all variables and the correlation between them.

Learning and Well-Being Goals

Figure 14 displays the average ratings of participants' learning and well-being goals, measured across 17 surveys administered at the beginning of each class. A sharp decline in participants' learning goals was observed in Survey 12, conducted on November 6, 2024. This drop was likely influenced by distractions related to the presidential election occurring at that time. Since these external factors were unrelated to the study, I decided to exclude the reported learning goals in Survey 12. Hierarchical regression models were conducted to examine how learning and well-being goals varied across surveys and in relation to exam weight. Figure 15 presents the model-predicted ratings of learning and well-being goals across the 17 surveys.

Learning Goals. A linear mixed-effects model with a random intercept for participants (1|ParticipantID) was conducted to examine if participants' learning goals changed across the 17

surveys administered throughout the semester[4]. Results showed that participants' learning goals significantly decreased as the semester progressed, $B = -0.03$, $SE = 0.01$, $p < .001$. In a subsequent step, we added exam weight to the model to investigate whether participants were more motivated to learn when anticipating an exam worth 25% of the final grade compared to an exam worth 5%. Results showed that the main effect of exam weight was not significant, $B = 0.01$, $SE = 0.05$, $p = .872$, indicating that participants' learning goals when expecting 25% exams ($M = 3.54$, $SD = 0.85$) were comparable to those when expecting 5% exams ($M = 3.47$, $SD = 0.90$). A likelihood ratio test conducted to compare the two models suggested that adding exam weight as a predictor of learning goals did not significantly improve the model fit, $\chi^2(1) = 0.03$, $p = .871$.

Well-Being Goals. A similar linear mixed-effects model was performed to examine how well-being goals changed over the course of the semester. The results showed that there was no significant variation in well-being goals across the semester, $B = -0.01$, $SE = 0.01$, $p = .099$. Next, exam weight was added as a predictor to assess its impact on well-being goals. Results indicated that exam weight was not a significant predictor, $B = -0.02$, $SE = 0.05$, $p = .712$, suggesting that participants' well-being goals were similar whether expecting 25% exams ($M = 3.70$, $SD = 0.78$) or 5% exams ($M = 3.72$, $SD = 0.84$). A likelihood ratio test conducted to

[4] Class format (in-person vs. remote) was initially included in the first step of the model predicting learning goals. However, it was not a significant predictor. Therefore, I removed class format from the model and retained only the survey variable to account for the overall declining trend in learning goals across the semester.

compare the two models suggested that adding exam weight as a predictor of well-being goals did not significantly improve the model fit, $\chi^2(1) = 0.14$, $p = .711$.

Exam Scores

Figure 16 displays boxplots of exam scores for the five unit exams, comparing performance by exam weight (5% vs. 25%). Although no significant effect of exam weight on learning and well-being goals was observed, I hypothesize that this may be due to the question items not being sensitive enough to capture the impact of exam weight on students' goals. I speculate that participants may be more motivated to study for exams weighted at 25%, thereby devoting more time and effort to preparing for these exams compared to those weighted at 5%. If this is the case, I would expect participants to perform better on the 25% exams.

To test this hypothesis, linear mixed-effects models with a random intercept for participants were conducted. Model 1, which included exam as the sole predictor, showed that exam scores significantly increased across the exams, $B = 2.93$, $SE = 0.48$, $p < .001$. Model 2 added exam weight as an additional predictor alongside exam. The results indicated that exam weight was also a significant predictor, with participants performing significantly better on exams weighted at 25% of the final grade ($M = 87.03$, $SD = 8.10$) compared to exams weighted at 5% ($M = 73.11$, $SD = 15.57$), $B = 10.07$, $SE = 0.94$, $p < .001$. A likelihood ratio test comparing Model 1 and Model 2 indicated that including exam weight significantly improved the fit, $\chi^2(1) = 115.21$, $p < .001$. This pattern held even when Exam 1 scores were excluded, as it was the only exam for which students did not receive a study guide. After removing Exam 1, participants still performed significantly better on 25%-weighted exams compared to 5%-weighted exams ($M = 79.69$, $SD = 12.81$), $B = 8.24$, $SE = 0.96$, $p < .001$.

Choice of Interval Practice

Figure 17 shows the percentage of participants who chose interval testing across the 17 surveys. At the start of the semester, about 50% of participants chose interval testing, but this percentage steadily declined over time. By Survey 6 (after Exam 2), the number of participants selecting retrieval practice dropped to around 20%, and it remained at this lower level until Survey 10. After Exam 3 and the memory lecture (from Survey 10 onward), the choice of retrieval practice increased slightly, reaching a peak of about 25% around Survey 13. However, after Exam 4 (starting with Survey 15), when classes shifted to remote learning, the percentage of participants choosing retrieval practice decreased again and stabilized at a lower level (around 10-15%).

Participants completed Survey 10 at the beginning of the class, just before receiving the memory lecture. Although this choice was made technically prior to exposure to the memory lecture, some students may have already reviewed the lecture slides, which were posted online before class. Given this possibility, the choice of practice format in Survey 10 was excluded from the analysis, as it may have been influenced by prior exposure to the lecture content.

Overview of the Analyses. A series of hierarchical linear regressions was conducted to examine the factors predicting participants' choice of interval practice. Table 11 presents a summary of the hierarchical regression analyses predicting interval practice choice in Experiment 4.

Model 1 predicted participants' choice of interval practice (0 = interval study, 1 = interval test) based on control variables, including participants' individual characteristics, class format (0 = remote and asynchronous; 1 = in-person and synchronous), and a quadratic term for Survey (to capture potential non-linear changes over time). Individual characteristics included frequency of

self-testing, test anxiety, unit exam scores, perceived strategy effectiveness, perceived anxiety and frustration, learning goals, and well-being goals. All predictors related to individual characteristics, along with the quadratic term for Survey, were mean-centered to provide a meaningful zero point and to facilitate interpretation. A random intercept for participants (1|Participant ID) was included to account for individual variability across repeated measures.

Model 2 extended Model 1 by including exam weight (0 = 5%, 1 = 25%) as an additional predictor. Model 3 further added memory lecture (0 = pre-exposure, 1 = post-exposure) as a predictor. Model 4 included an interaction term between learning goal and memory lecture to test whether the effect of learning goals on strategy choice differed depending on whether participants had been exposed to memory lectures emphasizing the benefits of retrieval practice for long-term memory.

Across these models, a total of 12 predictors were included, with three main variables of interest. A sensitivity analysis indicated that with 35 participants, the study had sufficient power to detect a large effect size of $f^2 = 0.373$.

Model 1: Control Variables. A logistic regression model was fitted to predict the choice of interval practice based on control variables. Results show that the quadratic term for survey was significant ($OR = 1.08$, 95% CI [1.05, 1.12], $p < .001$), suggesting a U-shaped pattern in strategy choice. Participants were less likely to select interval testing leading up to the mid-semester point (Survey 9), but their likelihood increased in the latter half of the semester. Class

format was also a significant predictor (*OR* = 76.69, 95% CI [5.44, 1080.67], $p = .001$)[5], indicating that participants were more likely to choose interval testing when the class was held in person and synchronously.

Among individual characteristics, perceived strategy effectiveness significantly predicted interval practice choice (*OR* = 2.68, 95% CI [1.66, 4.32], $p < .001$). Participants who viewed interval testing as more effective were more likely to choose it. Perceptions of anxiety and frustration approached significance (*OR* = 0.40, 95% CI [0.16, 1.01], $p = .053$), suggesting a trend where participants were less likely to choose interval testing if they associated it with higher anxiety and frustration. This effect became stronger and reached significance in Models 3 and 4, after accounting for the effect of memory lecture and its interaction with learning goals.

Well-being goals also significantly predicted interval practice choice (*OR* = 0.30, 95% CI [0.13, 0.70], $p < .01$). The more participants wanted to relax and enjoy the class, the less likely they were to choose interval testing. In contrast, learning goals were not a significant predictor (*OR* = 1.12, 95% CI [0.51, 2.48], $p = .772$). Figures 18 and 19 show the predicted probability of participants choosing interval testing based on different predictors: learning and well-being goals

[5] The odds ratio for class format should be interpreted with caution. When I ran a simpler model with class format as the sole predictor of interval practice choice—along with a random intercept for participants—class format remained a significant predictor (*OR* = 4.44, 95% CI [1.73, 11.39], $p = .002$), though the effect size was notably smaller. This suggests that including class format alongside other predictors may have inflated its estimated effect. However, given that class format consistently remained a significant predictor and there was no evidence of high multicollinearity among predictors, I chose to retain it in the analyses.

(Figure 18), and perceived strategy effectiveness along with perceived anxiety and frustration (Figure 19). Other predictors in the model were not significant (*ps* > .05).

Model 2: Adding Exam Weight. Results showed that exam weight was not a significant predictor (*OR* = 0.87, 95% CI [0.23, 3.30], *p* = .843). A likelihood ratio test comparing Model 2 to Model 1 demonstrated that the inclusion of exam weight did not significantly improve the model fit, $\chi^2(1) = 0.004$, $p = .953$.

Model 3: Adding Memory Lecture. Results showed that memory lecture was not a significant predictor participants' choice of interval practice (*OR* = 1.40, 95% CI [0.39, 4.96], *p* = .604). This model did not fit the data significantly better than Model 2, $\chi^2(1) = 0.30$, $p = .583$.

Model 4: Adding Interaction between Learning Goal and Memory Lecture. Results indicated that the interaction between learning goals and memory lecture was not significant (*OR* = 1.19, 95% CI [0.36, 3.95], *p* = .775). A likelihood ratio test comparing Model 3 to Model 2 revealed that adding the interaction term did not significantly improve the model fit, $\chi^2(1) = 0.010$, $p = .756$.

Study Outside of Class and Perceptions of Effectiveness, Anxiety and Frustration

Table 12 summarizes the study strategies participants reported using in the pre-class and mid-semester surveys.

Pre-class survey. A total of 32 participants completed the pre-class survey. Among them, 16 participants (50%) reported using retrieval practice techniques, including flashcards, recalling information from memory, and taking practice tests or quizzes. Seven participants (21.88%) reported using generation strategies, including summarizing, synthesizing and reorganizing materials, creating study guides, or explaining content to themselves or others. Eighteen participants (56.25%) indicated relying on more passive strategies, such as rereading or rewriting

notes and study materials (e.g., lecture slides, readings, homework, or textbook content). Other strategies included studying with peers (*n* = 4; 12.50%), watching online video lectures (*n* = 6; 18.75%), taking notes (*n* = 8; 25%), spacing out study sessions (*n* = 1; 3.12%), and using AI tools like ChatGPT (*n* = 1; 3.12%).

Mid-semester survey. There were 34 participants completing the mid-semester survey. Most of them reported spending 1-2 hours per week on this class outside of class meetings (*n* = 19; 55.88%). Eleven participants reported spending 3-4 hours per week (32.35%), two students reported spending more than 5 hours (5.88%), and two students reported spending less than 1 hour (5.88%).

Regarding study strategies, eight participants (*n* = 8; 23.53%) reported some forms of retrieval practice, such as using flashcards, recalling information from memory, answering end-of-chapter questions, or testing themselves on studied concepts. This was a smaller number of participants reporting the use of retrieval practice compared to the beginning of the semester. Nine participants (26.47%) mentioned generation strategies, including explaining materials in their own words, making study guides, or using concept maps. A larger number of participants (*n* = 26; 76.47%) relied on passive strategies, such as rereading and rewriting notes, and reviewing study guides, textbooks, and lecture slides. This represents a shift toward more passive study strategies compared to what was reported in the pre-class survey. Other strategies included studying with peers (*n* = 3; 8.82%), watching online video lectures (*n* = 2; 5.88%), taking notes (*n* = 8; 23.53%), and searching for studied concepts online (*n* = 1; 2.94%).

When asked which type of interval practice they used more often, 27 out of 34 students (79.41%) reported choosing interval study, while seven students (20.59%) said they used interval testing more frequently during the semester. Students who preferred interval study gave a range

of reasons. The most common reason, mentioned by 11 students (32.35%), was a lack of confidence—they did not feel ready to be tested. One student also said that they chose interval study because it felt easier. Ten students (29.41%) believed that interval study was more effective for their learning. Interestingly, seven students reported using interval study in a way that resembled testing. They first attempted to recall the information before checking the correct answers. One student said they preferred interval study because it matched their learning style. Another student believed interval testing was not helpful, while a third felt that only successful retrieval made testing beneficial. In addition, two students mentioned emotional reasons, describing interval study as a less stressful option.

Among the seven students who chose interval testing more often, three said it helped them learn better, while four valued it as a metacognitive tool for assessing what they knew and did not know. One student felt that both strategies were effective and reported using them interchangeably.

Regarding perceptions of strategy effectiveness, only 29.41% of participants believed that interval testing was more effective than interval study. However, the majority of participants (97.06%) felt that interval testing was more likely to induce anxiety and frustration.

Discussion

Experiment 4 aimed to explore how goals related to learning and well-being influence strategy choice in a real-world classroom setting, where mastering the material had direct consequences for college performance. This experiment allowed participants to select study strategies aligned with their personal and academic goals in the context of actual course assessments. The study manipulated learning goals by varying the weight of unit exams, assigning them either 5% or 25% of the final grade. To control for individual differences in

perceptions of strategy effectiveness, a series of memory lectures on the benefits of retrieval practice for long-term retention were provided midway through the semester. This design sought to examine how goal orientation, influenced by task importance, and perceptions of strategy effectiveness together shaped strategy choices.

The results showed no significant effect of exam weight on participants' reported learning and well-being goals or their choice of interval practice. However, participants performed better on the 25%-weighted exams than on the 5% exams, suggesting they may have studied more or used more effective strategies for higher-stakes exams. While we asked participants to report their general out-of-class study behavior, we did not specifically ask how their study approach varied by exam weight, limiting our ability to draw firm conclusions about how exam stakes influenced study strategies.

Participants reported using various forms of retrieval practice when studying outside of class. At the beginning of the semester, 50% of students reported using some form of retrieval practice when studying on their own, but by the middle of the semester, this dropped to just 24%. Additionally, there was an increase in reliance on passive strategies, such as rereading and rewriting materials, as the semester progressed. This shift in strategy use may be attributed to the growing workload participants faced, prompting them to adopt easier, more comfortable strategies as the demands of the semester intensified.

A key finding in this experiment is that only 29% of participants believed that interval testing was more effective than interval study, whereas 97% reported that interval testing was more likely to induce anxiety and frustration. This contrasts with prior experiments, where most participants recognized the effectiveness of retrieval practice. This discrepancy may stem from the nature of the materials studied in this experiment. Unlike Experiments 1-3, which involved

vocabulary tasks that required less cognitive effort and focused mainly on memorization, the materials here were more complex and required deeper understanding. Research suggests that the use of retrieval practice varies across learning domains (Tullis & Maddox, 2020). Learners are more likely to engage in retrieval practice for relatively easier materials, such as learning word pairs (Kornell & Son, 2009), but may underestimate its benefits for more cognitively demanding content, such as cognitive psychology, which requires deeper understanding and comprehension (Agarwal et al., 2008; Roediger & Karpicke, 2006). This difference in material complexity could explain why only 29% of participants perceived retrieval practice as effective in this context.

Another contributing factor may be that participants in this experiment had only one opportunity to complete the interval test questions, which did not reappear in subsequent interval sessions or on the unit exams. This design contrasts with Experiments 1–3, where participants completed three rounds of retrieval practice and were able to observe their improvement over time. Research shows that repeated retrieval enhances memory by reinforcing learning and improving recall accuracy across sessions (Higham et al., 2021). Without these repeated opportunities, participants in Experiment 4 may have had fewer chances to witness improvements resulting from interval testing, potentially contributing to the lower percentage of students who viewed it as effective.

The absence of significant effects of exam weight on strategy choice could be due to a lack of awareness of the benefits of retrieval practice and interval testing. This finding aligns with the results from Experiments 1B and 3, where goal orientation did not significantly impact strategy choice when students were not fully aware of the benefits of retrieval practice.

Importantly, participants' engagement in interval testing declined over time—from about half of the participants at the start of the semester to just 20% by the midpoint. There was a small

increase in its use after the memory lectures, reflected in a significant U-shaped pattern in choices over time. However, the memory lecture itself was not a significant predictor of interval practice choice. One possible reason is that the class format may have influenced the results. Participants were less likely to choose interval testing during the last three remote, asynchronous sessions—possibly because there was no instructor supervision and more distractions. This aligns with past research showing that students are less likely to use self-testing when studying on their own without supervision, even when they are informed of its mnemonic benefits (Simone et al., 2023).

Still, even during in-person classes, most participants chose interval study over interval testing. This low rate of engagement with interval testing may be due to participants' lack of confidence in producing correct answers, along with the misconception that retrieval practice only works when they can successfully recall the information. Research indicates that learners may fail to recognize the benefits of retrieval practice for long-term retention, but they do acknowledge its value as a diagnostic tool for identifying learning gaps (Hartwig & Dunlosky, 2012; Kornell & Bjork, 2007; Kornell & Son, 2009). As a result, learners are less likely to engage in retrieval practice when their confidence in retrieval success decreases. Additionally, some participants reported using interval study in a way that mimicked testing—by attempting to recall information before checking the answers—which may have reduced their motivation to adopt interval testing, even after receiving the memory lectures. Further, interval testing was also widely seen as more likely to induce anxiety and frustration. These emotional costs may have influenced students' strategy choices, particularly among those who prioritized emotional well-being.

Perceived strategy effectiveness and perceived anxiety and frustration significantly predicted participants' choice of practice format. Participants who viewed interval testing as more effective were more likely to choose it, whereas those who perceived it as anxiety- and frustration-inducing were less likely to do so. Well-being goals also significantly influenced strategy choice. Participants who expressed a stronger desire to relax and enjoy the class were less inclined to select interval testing. In contrast, learning goals did not significantly predict strategy choice. This was likely because most participants perceived interval study as more effective, and thus opted for it even when they were highly motivated to learn. These findings align with results from prior experiments, which showed that perceptions of strategy effectiveness, emotional cost, and well-being goals shape the use of retrieval practice. However, in this context, the low perceived effectiveness of retrieval practice appeared to diminish the influence of learning goals on participants' strategy choices.

In sum, Experiment 4 provides valuable insights into how goal orientation and perceptions of both the effectiveness and emotional cost of retrieval practice interact to influence strategy choice in a real-world classroom setting. The findings suggest that both learning and well-being factors influence students' engagement with retrieval practice. This experiment also highlights the complex nature of study behavior, which is guided not only by students' goals but also by their beliefs about study strategies, emotional experiences, and the broader learning context.

General Discussion

The primary goal of this study was to examine how learners' goals related to learning and emotional well-being influenced their allocation of study time to effortful learning strategies, such as retrieval practice. This study manipulate goals directly and provides evidence for the

causal effects of goal orientation on the regulation of study strategies. Across four experiments, the findings provided robust support for the effects of learning and emotional well-being goals on learners' engagement with effortful strategies like retrieval practice.

Effects of Goals on Strategy Choice in a Vocabulary Learning Task

Experiments 1 through 3 consistently found that participants in the well-being goal condition allocated significantly less time to retrieval practice compared to those in the no goal condition. This may be because, across all four experiments, over 80% of the participants in the no goal condition reported adopting a learning goal, even when no specific goal was provided. This pattern held true regardless of whether participants were allocating time for a hypothetical student (Experiments 1A, 1B, and 2) or for themselves (Experiment 3). The strong effects of goal orientation on strategy choice were observed consistently across Experiments 1-3, regardless of different time constraints (Experiment 1A) or the provision of research information on the learning benefits and emotional costs of retrieval practice (Experiment 2).

Interestingly, goal orientation had a stronger effect on strategy choice among participants who were aware of the impact of retrieval practice on learning and emotional well-being. Participants were more influenced by their assigned goal when they perceived retrieval practice as effective (Experiments 1B and 3) or were explicitly informed about the trade-off between learning gains and emotional costs through research information (Experiment 2). This interaction emerged in all experiments except Experiment 1A, where the varying time limit on practice trials may have disrupted the relationship between goal orientation and perceived strategy effectiveness.

Additionally, consistent with previous research (e.g., Kornell & Son, 2009), Experiments 1-3 showed that the majority of participants (more than 60%) recognized the benefits of retrieval

practice in learning word pairs. However, most participants also reported that retrieval practice was likely to induce anxiety and frustration (more than 80%), which further emphasizes the perceived emotional challenges that come with effortful learning. This aligns with findings from prior research which shows that while many learners acknowledge the effectiveness of retrieval practice, they also view it as emotionally taxing (Rea et al., 2022; Wang et al., 2023). Altogether, these findings suggest a trade-off between learning and emotional well-being when using retrieval practice, consistent with prior research (Wenzel, 2022). They also highlight the impacts of goals (learning and well-being goal) on the use of effortful learning strategies, aligning with SRL theories and frameworks (Boekaerts, 2011; Thomas et al., 2022).

Experiments 1 – 3 contribute to the expanding literature on motivated SRL by emphasizing the critical role of both cognitive and emotional factors in shaping students' engagement with effortful learning strategies. Specifically, the results align with the dual processing self-regulation model (Boekaerts, 2011; Boekaerts & Cascallar, 2006), which posits that learners choose between a well-being pathway or a mastery pathway and select a learning method that is consistent with their goals. In Experiments 1-3, we found that participants allocated time to retrieval practice differently depending on whether they prioritized learning or emotional well-being. Those who emphasized emotional well-being were less likely to engage in effortful strategies like retrieval practice compared to those who prioritized learning. This supports the dual processing model by demonstrating that learners adjust their strategies based on their primary goal (learning or well-being) associated with specific tasks.

Similarly, the Meta-Affect Metacognition (MAMC) framework (Thomas et al., 2022) emphasizes the dynamic interplay between cognitive and emotional motivations in shaping learning behaviors. MAMC proposes that learners monitor both their cognitive progress and

emotional experiences during learning, adjusting their strategies to achieve desired learning outcomes and emotional states. The findings of Experiments 1-3 provide empirical support for this framework by showing that learners not only track their cognitive progress but also evaluate their emotional responses to the learning process. Many participants acknowledged the effectiveness of retrieval practice but also perceived it as likely to provoke negative emotions, such as anxiety and frustration. As a result, when encouraged to prioritize well-being, participants adjusted their strategies, opting for the less demanding and emotionally taxing study practice rather than the more effortful retrieval practice. Importantly, learners' perceptions of strategy effectiveness appear to moderate the influence of goals on learning decisions. Goals impacted strategy choice only when participants recognized the benefits of retrieval practice. Altogether, these findings highlight the importance of incorporating both goal orientation and learners' perceptions of learning progress and emotional experience into models of motivated SRL, in line with the MAMC framework.

Together, these results support both the dual processing model and the MAMC framework, reinforcing the idea that learning is not solely a cognitive process but is profoundly influenced by emotional experiences. Experiments 1-3 provide empirical evidence for SRL models, highlighting the importance of incorporating both cognitive and emotional factors in understanding and predicting how learners engage with effortful learning strategies, as they navigate the trade-off between learning and emotional well-being.

Effects of Goals on Strategy Choice in a College Course

Experiment 4 extended the findings from Experiments 1-3 by examining how learning and well-being goals evolved over the course of a semester in a real classroom setting. The results showed that learners' desire to relax and enjoy the class (a well-being goal) remained

relatively consistent throughout the semester, while their desire to learn (a learning goal) declined over time. This decline may reflect the growing emotional and cognitive demands of coursework as the semester progressed, including responsibilities from this and other courses. Ongoing stress from frequent assessments and the cumulative nature of the material may have contributed to feelings of being overwhelmed or mentally fatigued, ultimately reducing students' motivation to engage with new content. Prior research has shown that, as students progress through a semester, they may experience fatigue, which can negatively affect motivation and engagement in effortful learning (Muenks et al., 2017). Interestingly, the stability of participants' well-being goals suggests that learning and well-being goals may have been considered independently. A decline in learning goals was not necessarily associated with an increase or decrease in well-being goals.

Exam weight (5% vs. 25% of the final grade) did not influence participants' choice of practice type for interval activities. However, students performed better on higher-weight exams (25%) compared to lower-weight exams (5%), suggesting that they likely devoted more time and effort to studying for higher-stakes exams. This pattern indicates that the non-significant effect of exam weight on strategy choice was unlikely due to a failure to influence participants' motivation to learn. Rather, it appears that most participants in Experiment 4 did not recognize the benefits of retrieval practice in the form of interval testing and instead believed that interval study was more helpful. As a result, the majority chose interval study to support their learning, regardless of exam weight.

The proportion of participants choosing interval testing throughout the semester was low, decreasing from 50% at the start of the semester to 20% by Survey 9. Although not statistically significant, the introduction of memory lectures in Experiment 4 slightly increased the use of

retrieval practice, with the proportion of participants choosing interval testing rising from approximately 20% to 30% following the lectures. However, despite the lectures highlighting the benefits of retrieval practice, it remained underutilized. This pattern aligns with previous research suggesting that awareness alone is not sufficient to alter actual study behaviors (Rea et al., 2022; Wang et al., 2023).

The underuse of retrieval practice was observed both in participants' strategy choices for interval activities during class and in the mid-semester survey, which assessed participants' study behaviors outside of class. The mid-semester survey revealed that only a small percentage of participants (23%) used retrieval practice outside of class.

Several factors may explain the underuse of retrieval practice in Experiment 4. One possible explanation is that fewer participants in Experiment 4 believed retrieval practice was beneficial for learning (29%) compared to Experiments 1-3, where the majority recognized its benefits (over 60%). Additionally, most participants acknowledged that retrieval practice was likely to induce anxiety and frustration (97%). Participants may have perceived negative feedback, such as retrieval failure or poor test performance, as a threat to their self-concept, which could have reduced their willingness to engage in self-testing (Cassady, 2004; Clark & Svinicki, 2015; Vaughn & Kornell, 2019). Consistent with previous research, emotional costs may have hindered engagement with retrieval practice, even when participants were aware of its effectiveness (Rea et al., 2022; Wang et al., 2023).

Further, participants may have believed that retrieval practice is only beneficial when they are successful in retrieving the content. Consequently, a lack of confidence in their retrieval abilities may have led to hesitation in using this strategy. Importantly, the one-time nature of the testing in Experiment 4 differed from the repeated testing used in Experiments 1-3. Without

experiencing improvement through repeated practice, participants in Experiment 4 may not have fully recognized the benefits of retrieval practice for their learning (Higham et al., 2022).

Finally, the classroom setting itself could have influenced participants' behaviors. Although performance on interval testing was graded by completion rather than accuracy, some participants may have still felt anxious about poor performance during interval testing or feared that it could negatively impact their grade or the instructor's evaluation. As a result, they may have opted for interval study over interval testing.

Consistent with findings from Experiments 1–3, participants' perceptions of strategy effectiveness significantly influenced their strategy choices—they were more likely to select interval testing when they believed it was more effective. Perceptions of anxiety and frustration also played a meaningful role. Participants were less likely to choose interval testing when they viewed it as more likely to induce anxiety and frustration. This pattern aligned with results from Experiment 2, where participants also provided more detailed ratings of the emotional impact of retrieval practice, rather than simply choosing between retrieval and study practice.

Regarding the effect of goal orientation, well-being goals—but not learning goals—appeared to influence strategy use. Participants were less likely to select interval testing when they reported a stronger desire to relax and enjoy the class. In contrast, the desire to learn new material did not significantly predict strategy choice. This may be because most participants perceived interval study as more effective for learning and chose it to support their goals, regardless of their level of learning motivation.

The findings of Experiment 4 offer further empirical support for the dual processing self-regulation model (Boekaerts, 2011; Boekaerts & Cascallar, 2006) and the MAMC framework (Thomas et al., 2022), both of which highlight the role of cognitive and emotional goals in

guiding students' study decisions. In line with these frameworks, well-being goals significantly influenced participants' strategy choices. Those who prioritized emotional well-being were more likely to avoid effortful strategies like retrieval practice. However, learning goals did not significantly shape strategy use in this classroom context. This may be due to limited metacognitive awareness. Many participants did not recognize the benefits of interval testing and instead favored interval study, regardless of how much they valued learning. These findings reinforce those of the earlier experiments by demonstrating that the influence of goals on strategy choice depends on learners' perceptions of strategy effectiveness. Taken together, the results underscore the importance of integrating learners' goals, perceived emotional experiences, and perceived strategy effectiveness into models of motivated SRL, as proposed by the MAMC framework.

Limitations

There are several limitations to consider in this study. In Experiments 1A, 1B, and 2, participants allocated study time between retrieval and study practice for a hypothetical student rather than regulating their own learning. This design allowed for tight experimental control and helped minimize the influence of individual differences in study habits on time allocation. However, it limits ecological validity and may not fully reflect how students study in real learning contexts. To address this limitation, Experiment 3 asked participants to allocate study time for their own learning material. However, similar to the earlier experiments, participants' choices in Experiment 3 were not enacted—they did not actually engage in further study based on their allocation of study time. As a result, the findings still reflect intended rather than actual strategy use, which may limit their generalizability to real-world academic settings.

These limitations were addressed in Experiment 4, in which participants regulated their own study strategies in the context of a real college course. I manipulated learning goals by varying the weight of unit exams, either 5% or 25% of the final grade. The impact of exam weight was reflected in participants' performance, with higher scores observed on the 25%-weighted exams. However, these differences should be interpreted with caution, as other variables may have influenced exam performance. Specifically, the absence of a study guide for Exam 1 could have contributed to its lower performance. We began providing a study guide only after observing poor performance on the first exam. Additionally, Exam 5 was optional and only taken by students dissatisfied with their Exam 1 scores. This likely means that only low performers from Exam 1 opted for Exam 5, which may have resulted in lower scores for both Exams 1 and 5 compared to the other exams. Given the real-world nature of the classroom setting, this flexibility was inevitable, particularly since the exams were tied to students' GPAs. Furthermore, there may have been an order effect, with the sequence of exams influencing performance (e.g., Exam 1: 5%, Exam 2: 25%, Exam 3: 5%, Exam 4: 25%, Exam 5: 5%). Participants could have improved over time as they became more familiar with the exam content and study strategies. However, regression analyses showed that the effect of exam weight on exam scores remained significant even after accounting for natural improvement over time and excluding Exam 1 scores.

Additionally, exam weight did not significantly predict participants' choices for interval practice. This may be due to limitations in the design that prevented us from capturing how students adjusted their study behavior outside of class. While we measured general study habits, we did not assess whether participants studied differently for low- versus high-weighted exams, which could have influenced their strategy choices in ways we could not observe.

Similarly, the memory lecture on the mnemonic benefits of retrieval practice did not significantly impact participants' strategy selection, although we observed a slight increase in participants' engagement with interval testing after the memory lecture. One possible explanation is that the final three class sessions were conducted remotely and asynchronously. Without instructor supervision, students may have been less engaged or attentive during these sessions, which could have influenced their choice of interval practice. As a result, even after being explicitly informed of the benefits of retrieval practice, fewer participants chose interval testing during the remote and asynchronous sessions. Future research should revisit the effects of explicit instruction and task importance using a more controlled and comprehensive design that allows for clearer measurement of these influences on study behavior.

Practical Implications

While previous research has primarily attributed the underuse of retrieval practice to a lack of metacognitive awareness regarding its benefits (Rivers, 2021), the present study highlights the significant role of emotional costs—such as anxiety and frustration—and well-being goals in shaping strategy use. Learners may be hesitant to engage in retrieval practice when prioritizing emotional well-being due to the strategy's challenging nature. This reluctance persists even when learners are aware of the cognitive benefits of retrieval practice.

These findings carry important implications for educational practices and interventions aimed at promoting effective study strategies. Previous interventions have largely focused on enhancing learners' understanding of the benefits of retrieval practice and fostering learning goals, often overlooking the emotional impacts and the role of well-being goals. For example, the Knowledge, Belief, Commitment, and Planning (KBCP) framework proposed by McDaniel and Einstein (2020) emphasizes learners' knowledge of a strategy, their belief in its effectiveness,

their commitment to using it, and their planning for its implementation in learning. While these factors are essential, they may not address the emotional costs associated with effortful strategies.

Based on the findings of the present study, interventions designed to promote retrieval practice should also consider students' emotional goals and work to minimize the emotional costs associated with these strategies. Additionally, interventions could benefit from integrating affective research that helps shift learners' perception of emotional costs. For example, when students view negative emotions like anxiety as part of the learning process and an indicator of growth, they may come to appreciate these challenges as necessary for academic success (Yeager, et al., 2016). Reinterpreting such emotional experiences as essential components of learning could increase engagement with effortful strategies, ultimately improving learning outcomes.

Table 1

Perceptions of Strategy Effectiveness, Anxiety, and Frustration Across Goal Conditions and Practice Trial Time Limits in Experiment 1A

Time Limit	Goal Condition	*N*	Retrieval Practice as more effective	Retrieval Practice as more anxiety-inducing
5 seconds per retrieval trial	No Goal	47	61.70%	93.62%
	Learning	41	58.54%	87.80%
	Well-being	45	57.78%	88.89%
7 seconds per retrieval trial	No Goal	34	70.59%	82.35%
	Learning	41	63.41%	85.37%
	Well-being	40	70.00%	82.50%

Table 2

Summary of Hierarchical Regression Analyses for Predicting the Time Allocated to Retrieval Practice in Experiment 1A

	Model 1		Model 2		Model 3		Model 4	
Variable	*B*	*SE*	*B*	*SE*	*B*	*SE*	*B*	*SE*
Intercept	40.22***	4.05	42.96***	4.31	42.79***	4.38	40.41***	4.78
Retrieval practice performance (round 3)	8.39	4.30	8.48*	4.19	8.31	4.26	8.33	4.27
Retrieval practice as more effective	21.67***	2.60	21.67***	2.53	21.68***	2.54	25.80***	4.10
Retrieval practice as more anxiety-inducing	-5.77	3.28	-6.06	3.20	-5.99	3.22	-6.34	3.24
Learning goal			0.91	2.63	0.86	2.64	4.48	4.35
Well-being goal			-8.35**	2.60	-8.37**	2.61	-4.15	4.38
7 seconds per retrieval trial					0.52	2.19	0.50	2.19
Retrieval practice as more effective × Learning goal							-5.64	5.47
Retrieval practice as more effective × Well-being goal							-6.54	5.47
R^2	0.34		0.38		0.38		0.38	
F			7.78***		0.06		0.83	

Note. For perception variables, 'study practice' was the reference level. For goal condition, the reference level was 'no goal'. For time limit, the reference level was '5 seconds per retrieval trial'.
** p < .05. ** p < .01. *** p < .001.*

Table 3

Perceptions of Strategy Effectiveness, Anxiety, and Frustration Across Goal Conditions in Experiments 1B

Goal	*N*	Retrieval Practice as more effective	Retrieval Practice as more anxiety-inducing
No goal	84	71.43%	84.52%
Learning	128	63.28%	87.50%
Well-being	84	75.00%	89.29%

Table 4

Summary of Hierarchical Regression Analyses for Predicting the Time Allocated to Retrieval Practice in Experiment 1B

	Model 1		Model 2		Model 3	
Variable	*B*	*SE*	*B*	*SE*	*B*	*SE*
Intercept	32.30***	4.15	35.69***	4.17	35.04***	4.80
Retrieval practice performance (round 3)	5.26	4.33	6.65	4.07	7.21	4.04
Retrieval practice as more effective	22.86***	2.59	23.64***	2.44	24.39***	4.42
Retrieval practice as more anxiety-inducing	3.65	3.36	4.49	3.15	4.18	3.14
Learning goal			-2.24	2.54	-4.52	4.51
Well-being goal			-16.44***	2.79	-8.63	5.37
Retrieval practice as more effective × Learning goal					3.71	5.44
Retrieval practice as more effective × Well-being goal					-10.47	6.27
R^2	0.26		0.35		0.37	
F			21.23***		3.25*	

Note. For perception variables, 'study practice' was the reference level. For goal condition, the reference level was 'no goal'.
* $p < .05$. ** $p < .01$. *** $p < .001$.

Table 5

Perceptions of Strategy Effectiveness, Anxiety, and Frustration Across Goal and Research Provision Prior to Exposure to Research Information in Experiment 2

Research	Goal	*N*	Retrieval Practice as more effective	Retrieval Practice as more anxiety-inducing
Absent	Learning	68	57.35%	67.65%
	Well-Being	80	58.75%	63.75%
Provided	Learning	73	68.49%	82.19%
	Well-Being	72	54.17%	66.67%

Table 6

Summary of Hierarchical Regression Analyses for Predicting the Time Allocated to Retrieval Practice in Experiment 2

	Model 1		Model 2		Model 3		Model 4	
Variable	*B*	*SE*	*B*	*SE*	*B*	*SE*	*B*	*SE*
Intercept	48.28***	1.43	61.84***	1.68	63.54***	1.97	59.16***	2.18
(Prior ratings) Perceived effectiveness	2.70***	0.38	2.34***	0.32	2.33***	0.32	2.24***	0.31
(Prior ratings) Perceived anxiety and frustration	-0.24	0.44	-0.85*	0.37	-0.81*	0.37	-0.93*	0.36
Well-being goal			-24.15***	2.11	-24.28***	2.11	-15.81***	2.86
Research provided					-3.40	2.09	5.60	2.93
Well-being goal × Research provided							-17.31***	4.07
R^2	0.16		0.42		0.43		0.46	
F			130.59***		2.65		18.10***	

Note. For perception variables, ratings for those who chose 'retrieval practice' remained unchanged (1 to 5), while ratings for those who chose 'study practice' were reversed (-1 to -5). For goal condition, 'learning goal' was the reference level. For the research information variable, 'research absent' was the reference level.
* $p < .05$. ** $p < .01$. *** $p < .001$.

Table 7

Perceptions of Strategy Effectiveness, Anxiety, and Frustration Across Goal Conditions in Experiment 3

Goal	*N*	Retrieval Practice as more effective	Retrieval Practice as more anxiety-inducing
No goal	93	65.59%	86.02%
Learning	95	61.05%	80.00%
Well-being	96	64.58%	81.25%

Table 8

Summary of Hierarchical Regression Analyses for Predicting the Time Allocated to Retrieval Practice in Experiment 3

	Model 1		Model 2		Model 3	
Variable	*B*	*SE*	*B*	*SE*	*B*	*SE*
Intercept	31.64***	4.10	35.50***	4.49	34.56***	5.01
Retrieval practice performance (round 3)	17.67***	4.48	17.49***	4.38	17.73***	4.36
Retrieval practice as more effective	18.27***	2.69	18.40***	2.62	19.97***	4.49
Retrieval practice as more anxiety-inducing	-4.09	3.20	-4.36	3.13	-4.67	3.11
Learning goal			-0.45	2.91	-3.52	4.78
Well-being goal			-10.16***	2.90	-3.60	4.86
Retrieval practice as more effective × Learning goal					5.13	6.02
Retrieval practice as more effective × Well-being goal					-10.14	6.03
R^2	0.25		0.29		0.31	
F			7.94***		3.44*	

Note. For perception variables, 'study practice' was the reference level. For goal condition, the reference level was 'no goal'.
* $p < .05$. ** $p < .01$. *** $p < .001$.

Table 9

Overview of Course Schedule and Survey Timeline in Experiment 4

Week	Monday	Wednesday
0	9/2: Holiday (No Class)	9/4: Standard Lecture Class (No Survey; Interval Study)
1	9/9: Standard Lecture Class (No Survey; Interval Testing & Interval Study)	9/11: Standard Lecture Class (Survey 1)
2	9/16: Standard Lecture Class (No Survey; practice type based on Survey 1 choice)	**9/18: Unit Exam 1 (5% weight)**
3	9/23: Standard Lecture Class (Survey 2)	9/25: Standard Lecture Class (Survey 3)
4	9/30: Standard Lecture Class (Survey 4)	10/2: Review Class (Survey 5)
5	**10/7: Unit Exam 2 (25% weight)**	10/9: Standard Lecture Class (Survey 6)
6	10/14: Holiday (No Class)	10/16: Standard Lecture Class (Survey 7)
7	10/21: Standard Lecture Class (Survey 8)	10/23: Review Class (Survey 9)
8	**10/28: Unit Exam 3 (5% weight)**	10/30: Standard Lecture Class (Survey 10) *** Memory Lecture Started**
9	11/4: Standard Lecture Class (Survey 11)	11/6: Standard Lecture Class (Survey 12)
10	11/11: Holiday; rescheduled to 11/12 (Tuesday): Standard Lecture Class (Survey 13)	11/13: Review Class (Survey 14)
11	**11/18: Unit Exam 4 (25% weight)**	11/20: Standard Lecture Class (Survey 15) *** Remote Class**
12	11/25: Standard Lecture Class (Survey 16) *** Remote Class**	11/27: Holiday (No Class)
13	12/2: Standard Lecture Class (Survey 17) *** Remote Class**	12/4: Course Review & Experiment Explanation
14	**12/9: Unit Exam 5 (5% weight, make-up exam for Exam 1)**	

Note. In surveys 1 through 17, participants reported their learning and well-being goals, along with their choice of practice type (interval testing or interval study). Classes were conducted in person unless otherwise indicated as a Remote Class.

Table 10

The descriptive statistics for all variables in Experiment 4 and the correlation between them

	Variable	*Mean (SD)*	Min-Max	1	2	3	4	5	6	7	8	9	10
1.	Self-Testing Frequency	2.88 (0.82)	1 - 4										
2.	Test Anxiety	43.12 (13.35)	20 - 72	-0.06									
3.	Exam Score	82.86 (12.01)	25 - 100	-0.09	-0.09								
4.	Perceived Effectiveness	-1.00 (2.88)	-4 - 4	0.04	-0.25***	0.19***							
5.	Perceived Anxiety and Frustration	3.24 (1.50)	-2 - 5	0.1*	0.40***	-0.03	-0.36***						
6.	Learning Goal	3.51 (0.88)	1 - 5	0.05	0.04	-0.23***	-0.15**	0.18***					
7.	Well-Being Goal	3.71 (0.81)	2 - 5	0.21***	0.13**	-0.08	-0.13**	0.29***	0.28***				
8.	Class Format	-	0 = Remote; 1 = In-person	0	0	0.09*	0	0	0.11*	0.005			
9.	Choice of interval practice	-	0 = Interval Study; 1= Interval Test	-0.04	-0.12**	0.11*	0.62***	-0.37***	-0.15**	-0.21**	0.09*		
10.	Exam weight	-	0 = 5% ; 1 = 25%	0	0	0.42***	0	0	0.04*	-0.01*	0.49***	0.05	
11.	Memory lecture	-	0 = Pre-Exposure; 1 = Post-Exposure	0	0	0.27***	0	0	-0.13**	-0.03	0.06	0.13**	0.54***

Note. * $p < .05$. ** $p < .01$. *** $p < .001$.

Table 11

Summary of Hierarchical Regression Analyses for Predicting Choice of Interval Practice in Experiment 4

	Model 1		Model 2		Model 3		Model 4	
Variable	*OR*	95% CI	*OR*	95% CI	*OR*	95% CI	*OR*	95% CI
Intercept	0.0002***	[< .00001, 0.006]	0.0002***	[< .00001, 0.006]	0.0002***	[< .00001, 0.006]	0.0001***	[< .00001, 0.006]
Survey (quadratic trend)	1.08***	[1.05, 1.12]	1.08***	[1.05, 1.12]	1.08***	[1.05, 1.12]	1.08***	[1.05, 1.12]
Self-testing frequency	0.84	[0.22, 3.11]	0.80	[0.22, 2.94]	0.84	[0.23, 3.01]	0.82	[0.23, 2.97]
Test anxiety	1.07	[0.98, 1.18]	1.07	[0.98, 1.17]	1.07	[0.98, 1.17]	1.07	[0.98, 1.17]
Exam score	1.01	[0.97, 1.06]	1.02	[0.95, 1.09]	1.02	[0.95, 1.09]	1.02	[0.95, 1.09]
Perceived effectiveness	2.68***	[1.66, 4.32]	2.62***	[1.64, 4.20]	2.58***	[1.62, 4.09]	2.56***	[1.62, 4.06]
Perceived anxiety and frustration	0.40	[0.16, 1.01]	0.41	[0.17, 1.01]	0.40*	[0.16, 0.98]	0.40*	[0.16, 0.99]
Learning goal	1.12	[0.51, 2.48]	1.13	[0.51, 2.48]	1.13	[0.52, 2.47]	1.04	[0.39, 2.79]
Well-being goal	0.30**	[0.13, 0.70]	0.30**	[0.13, 0.70]	0.30**	[0.13, 0.71]	0.32**	[0.14, 0.74]
Class format	76.69**	[5.44, 1080.67]	78.37**	[4.99, 1230.51]	113.79**	[4.78, 2709.18]	142.22**	[5.55, 3645.24]
Exam weight			0.87	[0.23, 3.30]	0.77	[0.18, 3.36]	0.72	[0.16, 3.17]
Memory lecture					1.40	[0.39, 4.96]	1.50	[0.42, 5.35]
Learning goal × Memory lecture							1.19	[0.36, 3.95]
AIC	194.20		196.20		197.90		199.80	
χ^2			0.004		0.30		0.10	

Note. OR and 95% CI refer to the odds ratio and the 95% confidence interval, respectively.
* $p < .05$. ** $p < .01$. *** $p < .001$.

Table 12

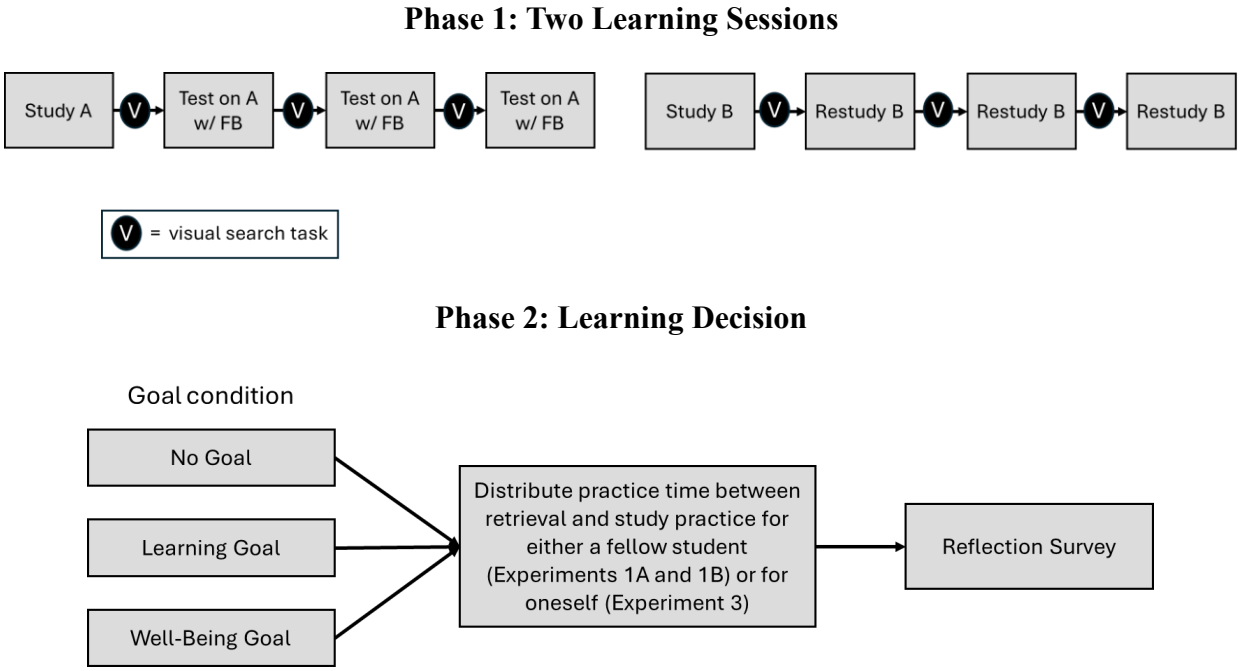
Reported Study Strategies in the Pre-Class and Mid-Semester Surveys in Experiment 4

Study Strategy	Pre-Class Survey (n = 32)	Mid-Semester Survey (n = 34)
Retrieval Practice (e.g., flashcards, self-testing)	16 (50%)	8 (23.53%)
Generation Strategies (e.g., summarizing, study guides)	7 (21.88%)	9 (26.47%)
Passive Strategies (e.g., rereading, reviewing notes)	18 (56.25%)	26 (76.47%)
Studying with Peers	4 (12.50%)	3 (8.82%)
Watching Online Video Lectures	6 (18.75%)	2 (5.88%)
Taking Notes	8 (25%)	8 (23.53%)
Spacing Study Sessions	1 (3.12%)	-
Using AI Tools (e.g., ChatGPT)	1 (3.12%)	-
Searching for Concepts Online	-	1 (2.94%)

Note. Values represent the number of students who reported using each type of study strategy. Percentages (in parentheses) were calculated out of the total number of respondents.

Figure 1

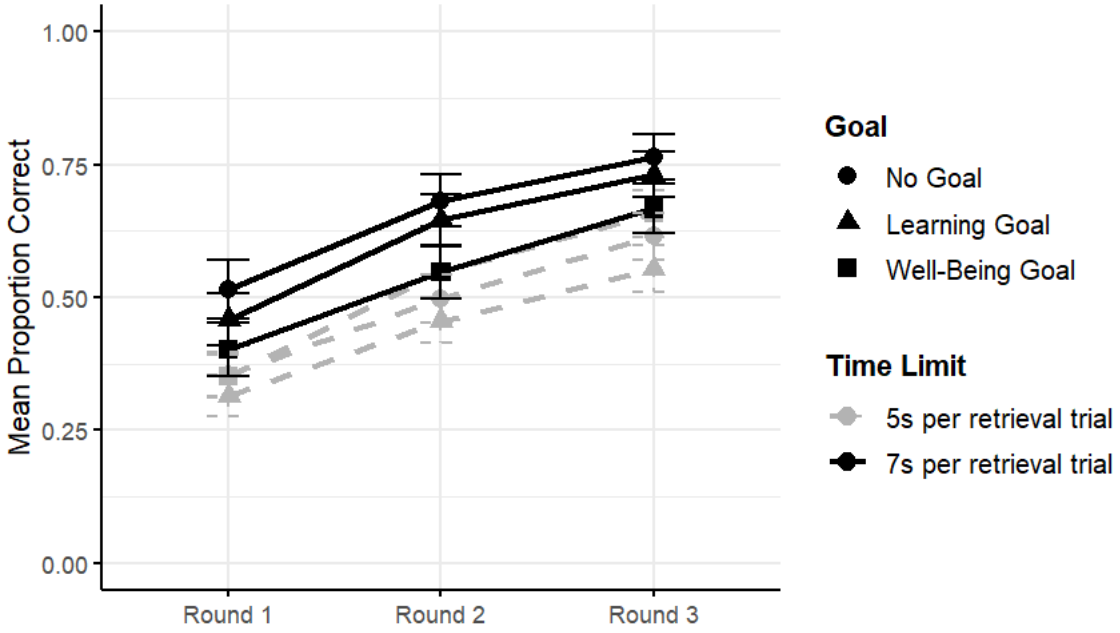
Overall Procedure of Experiments 1A, 1B, and 3



Note. The assignment of materials to each session (A or B) and the order of study strategies (retrieval practice or study practice) were counterbalanced. "Test on A w/ FB" refers to retrieval practice with feedback, and "Restudy B" refers to study practice. The key difference between Experiments 1A and 1B was the time limit per practice trial. In Experiment 1A, the time limit alternated between 5 and 7 seconds for retrieval practice and study practice. In contrast, Experiment 1B used a consistent time limit of 7 seconds per trial across both strategies.

Figure 2

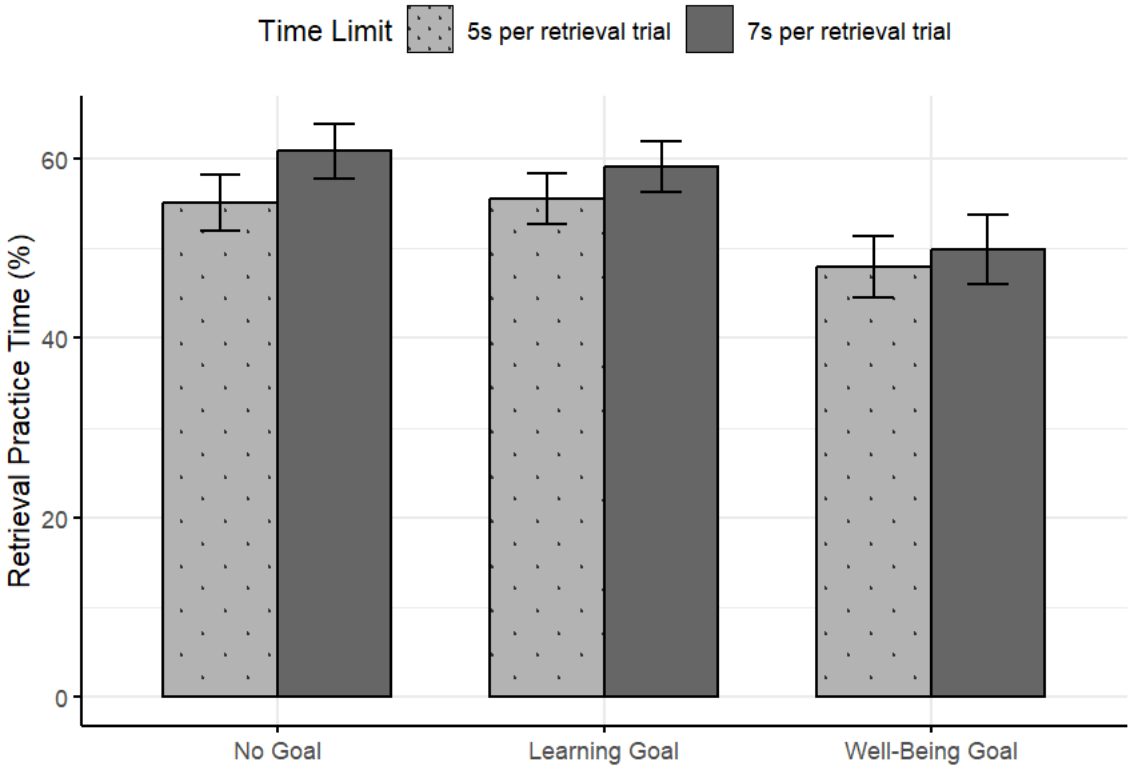
Retrieval Practice Performance Across Rounds, Goal Conditions, and Practice Trial Time Limit in Experiment 1A



Note. Error bars represent standard errors (SE). The labels '5s' and '7s' denote 5 seconds and 7 seconds, respectively.

Figure 3

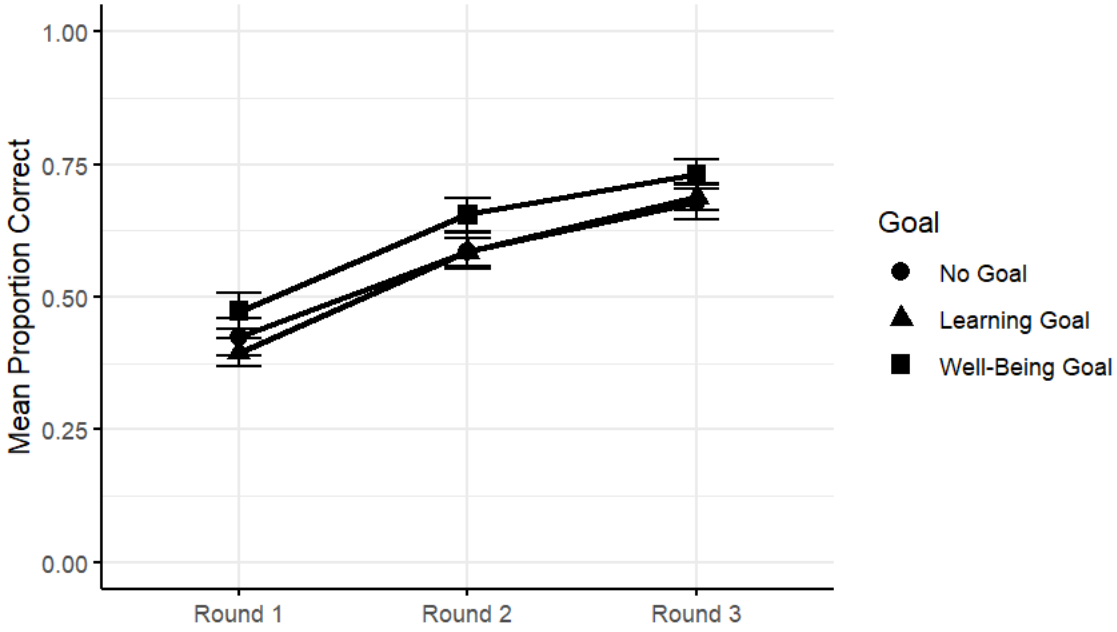
Time Allocated to Retrieval Practice Across Goal Conditions and Practice Trial Time Limit in Experiment 1A



Note. Error bars represent standard errors (SE). The labels '5s' and '7s' denote 5 seconds and 7 seconds, respectively.

Figure 4

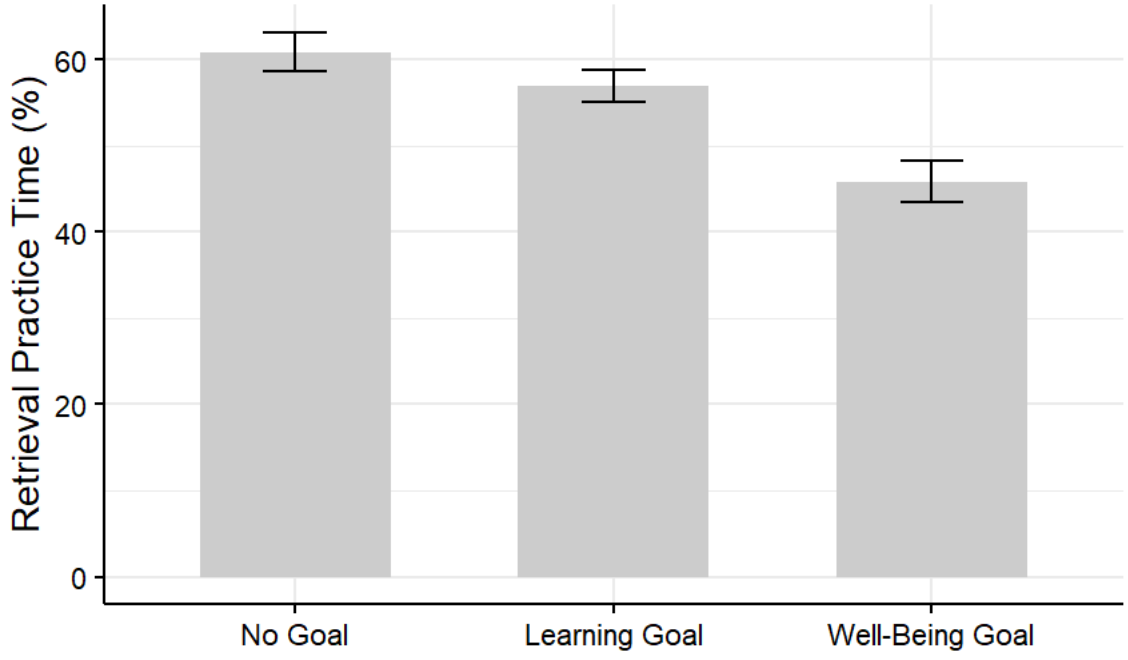
Retrieval Practice Performance Across Rounds and Goal Conditions in Experiment 1B



Note. Error bars indicate standard errors (SE).

Figure 5

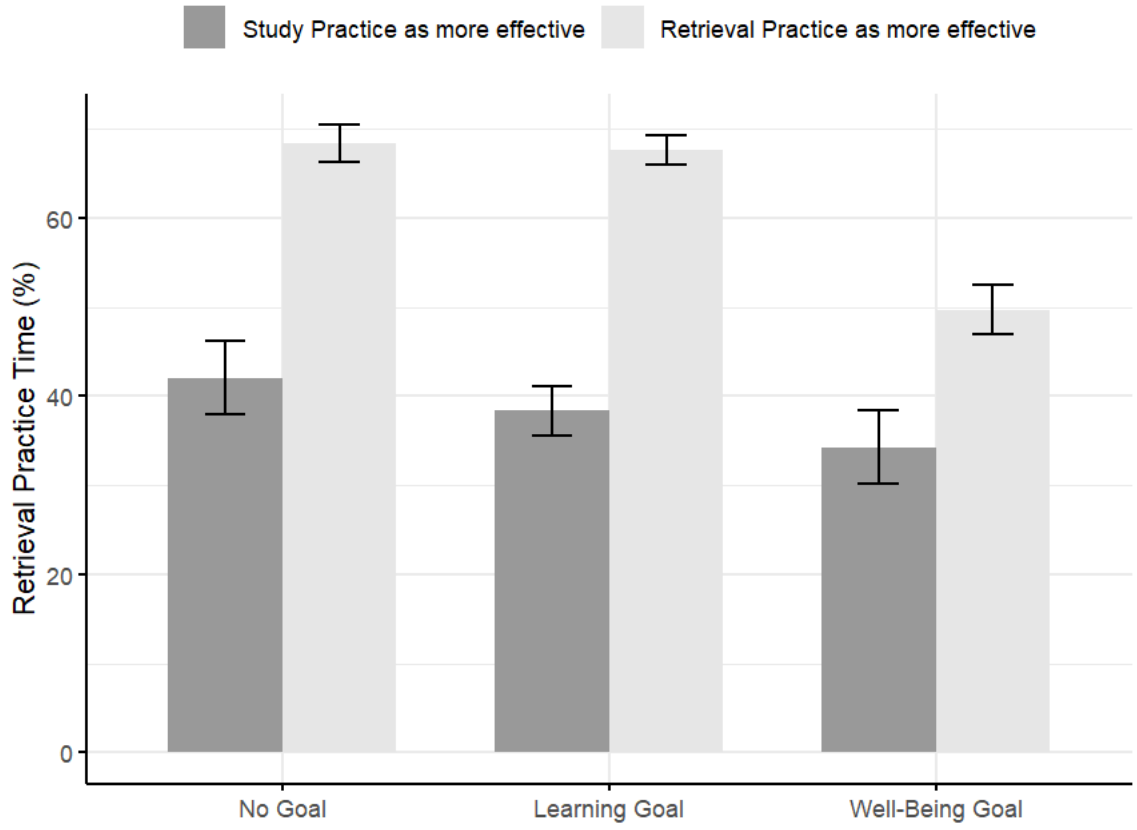
Time Allocated to Retrieval Practice Across Goal Conditions in Experiment 1B



Note. Error bars indicate standard errors (SE).

Figure 6

Time Allocated to Retrieval Practice Across Goal Conditions and Perception of Strategy Effectiveness in Experiment 1B



Note. Error bars indicate standard errors (SE).

Figure 7

Overall Procedure of Experiment 2

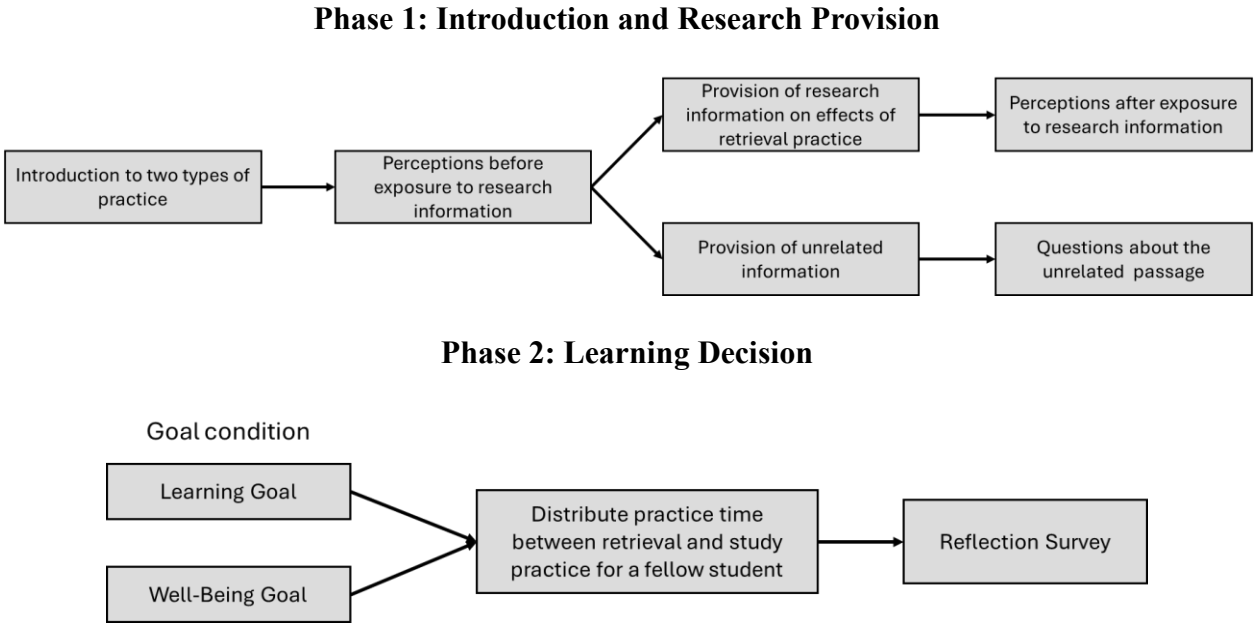
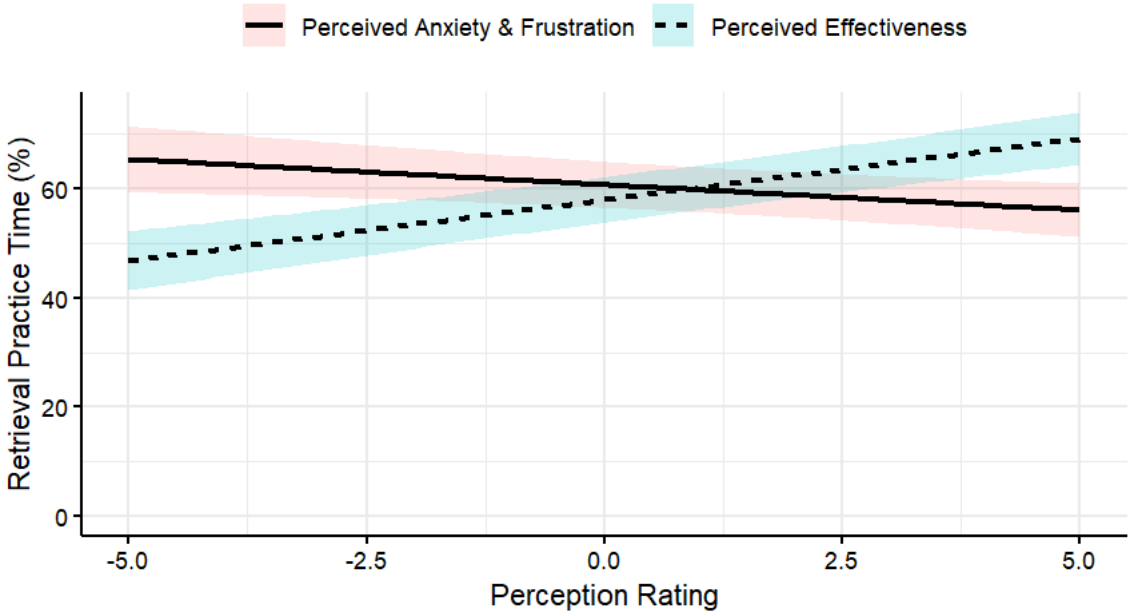


Figure 8

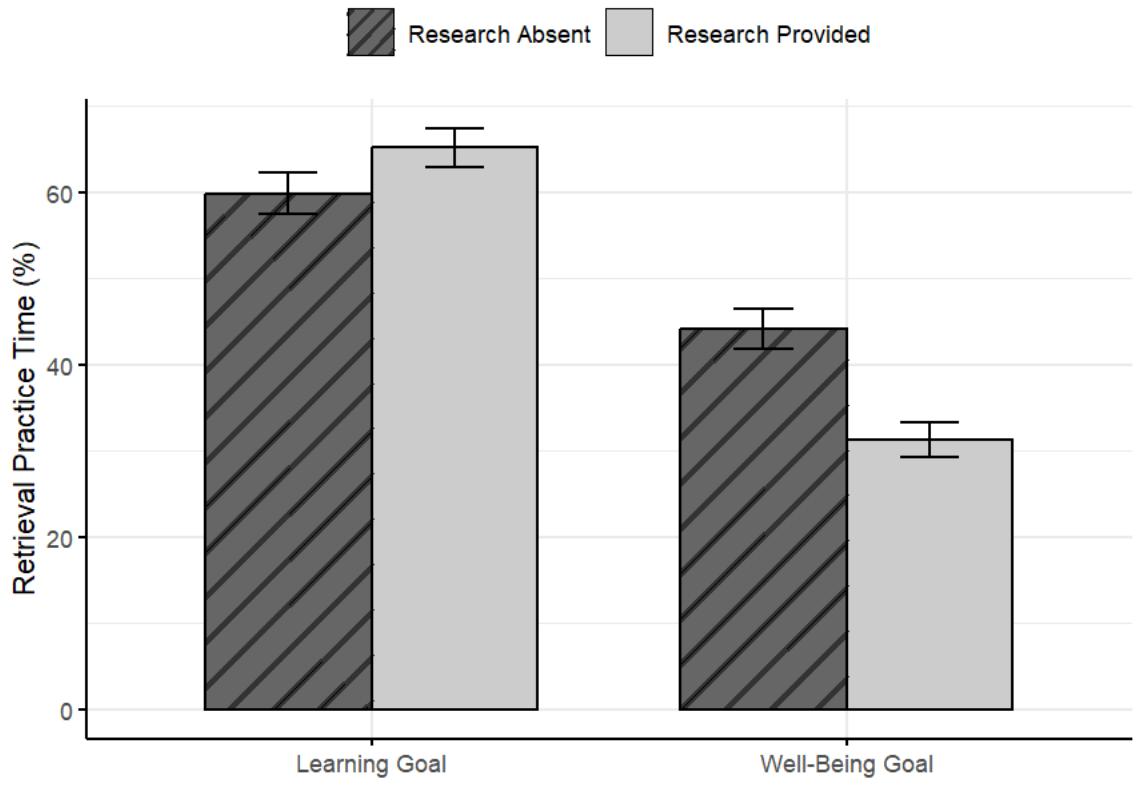
Predicted Time Allocated to Retrieval Practice as a Function of Prior Perceptions of Effectiveness, Anxiety, and Frustration in Experiment 2



Note. Shaded areas represent 95% confidence intervals around the predicted values.

Figure 9

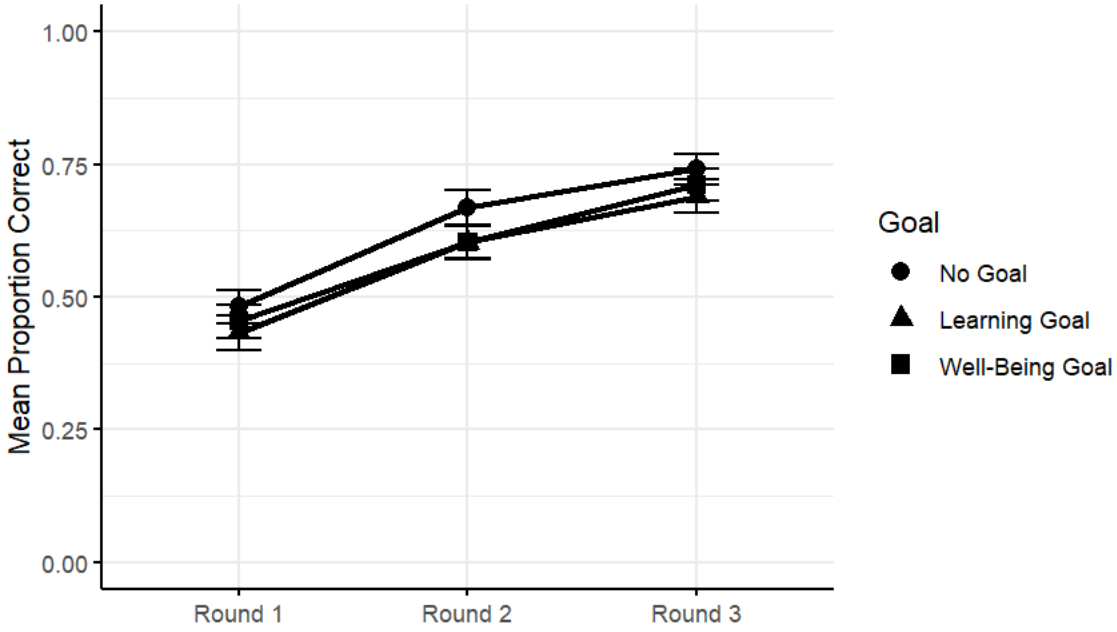
Time Allocated to Retrieval Practice Across Goal and Research Provision Conditions in Experiment 2



Note. Error bars indicate standard errors (SE).

Figure 10

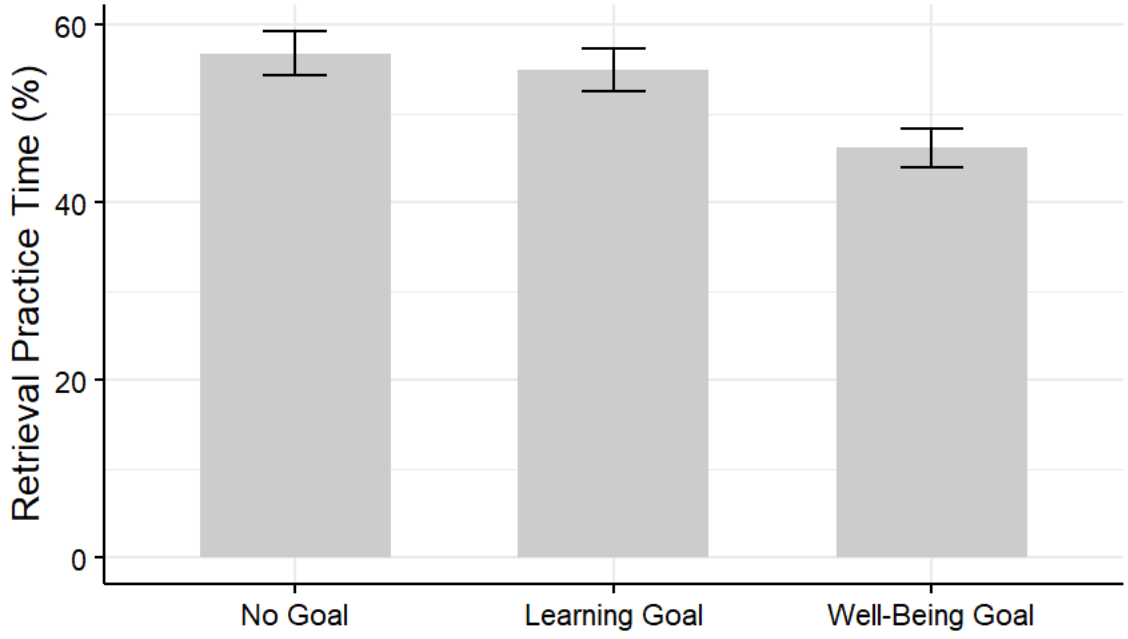
Retrieval Practice Performance Across Rounds and Goal Conditions in Experiment 3



Note. Error bars indicate standard errors (SE).

Figure 11

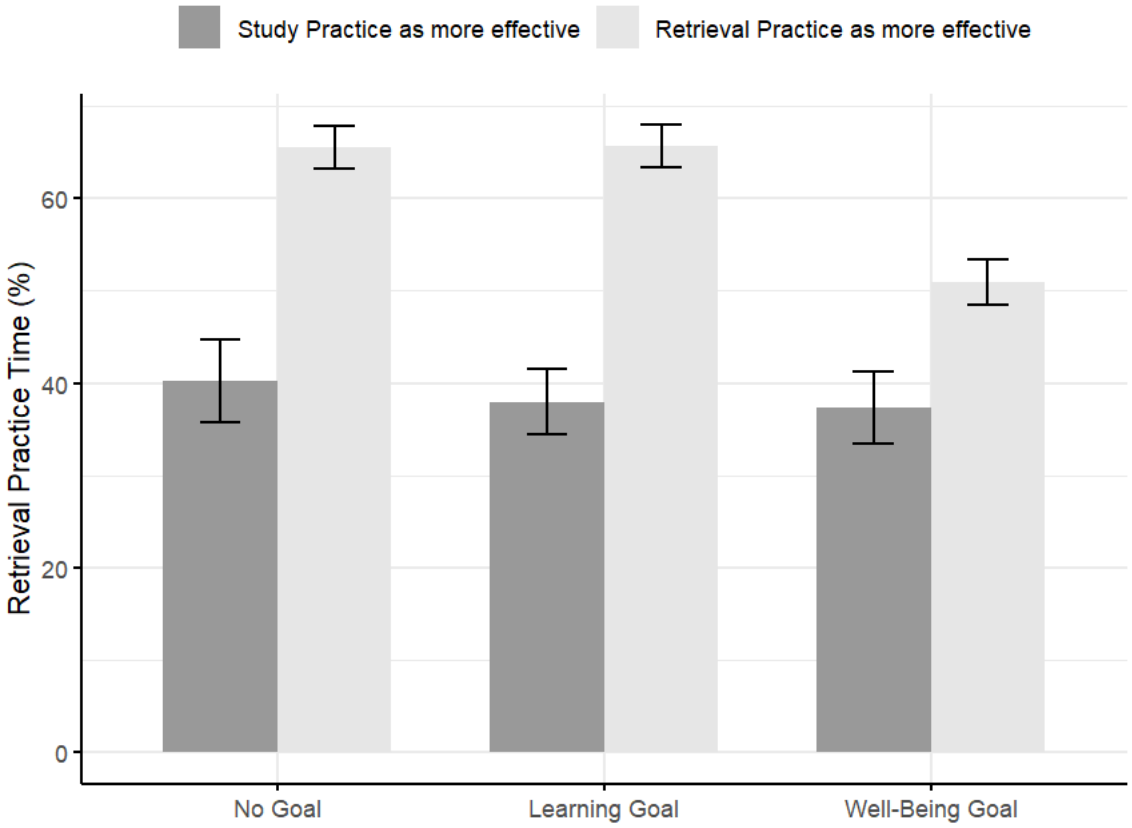
Time Allocated to Retrieval Practice Across Goal Conditions in Experiment 3



Note. Error bars indicate standard errors (SE).

Figure 12

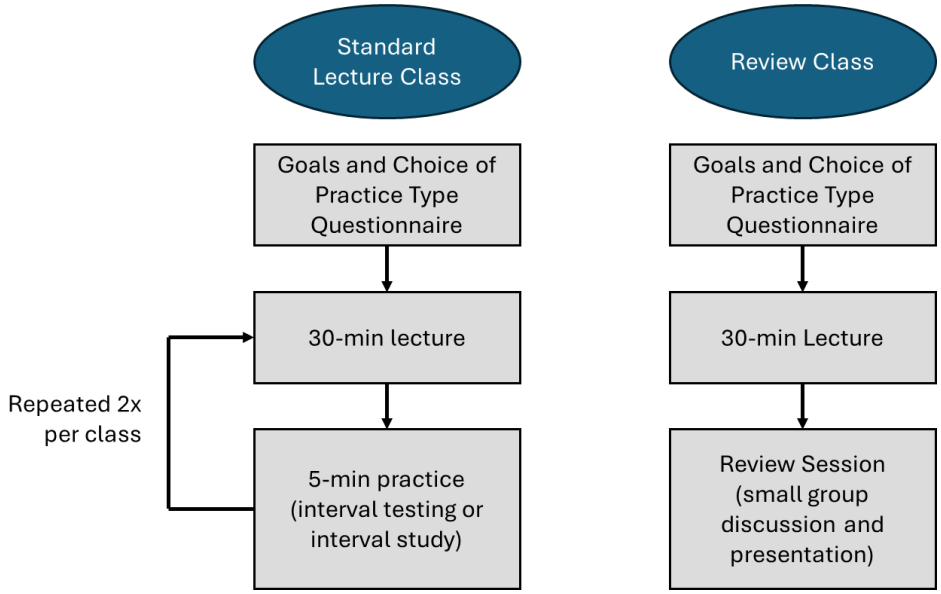
Time Allocated to Retrieval Practice Across Goal Conditions and Perception of Strategy Effectiveness in Experiment 3



Note. Error bars indicate standard errors (SE).

Figure 13

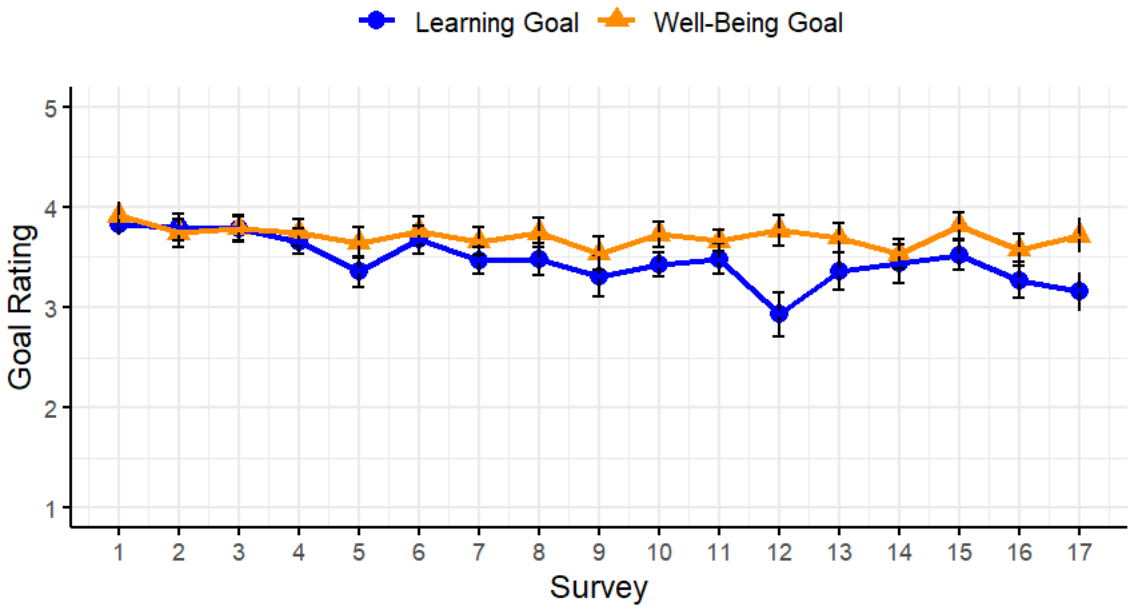
Class Structure for Standard Lecture Class and Review Class in Experiment 4



Note. Occasionally, due to time constraints, participants were unable to complete the second interval practice activity during Standard Lecture Classes. The structure of Standard Lecture Classes applied to both in-person and remote sessions, while Review Classes were always held in person.

Figure 14

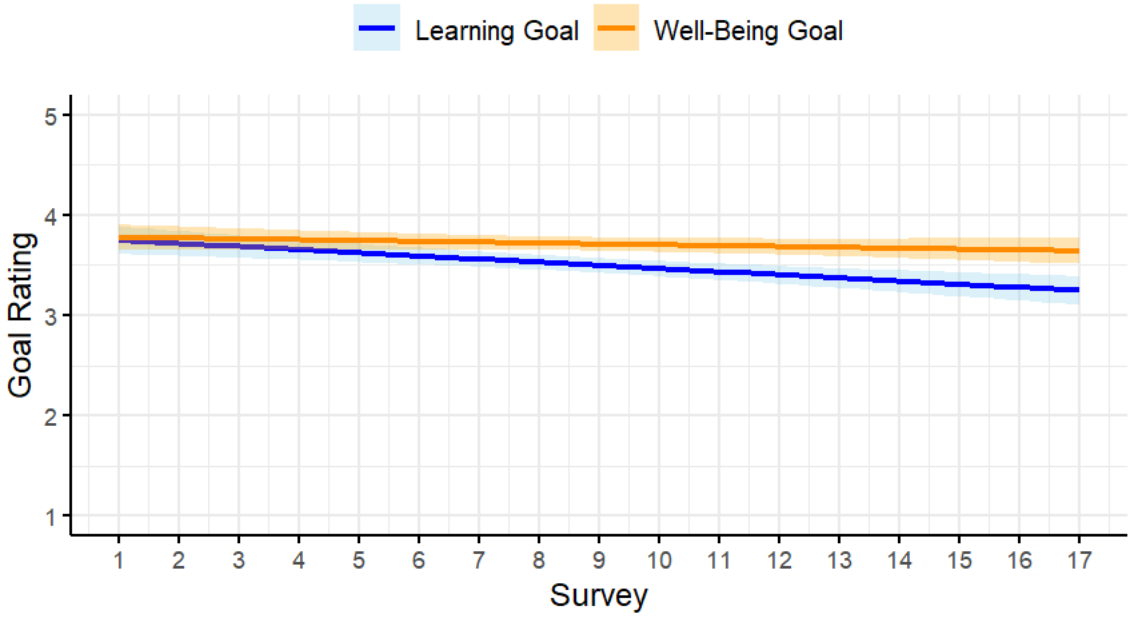
Learning and Well-Being Goal Ratings across Surveys in Experiment 4



Note. Error bars indicate standard errors (SE).

Figure 15

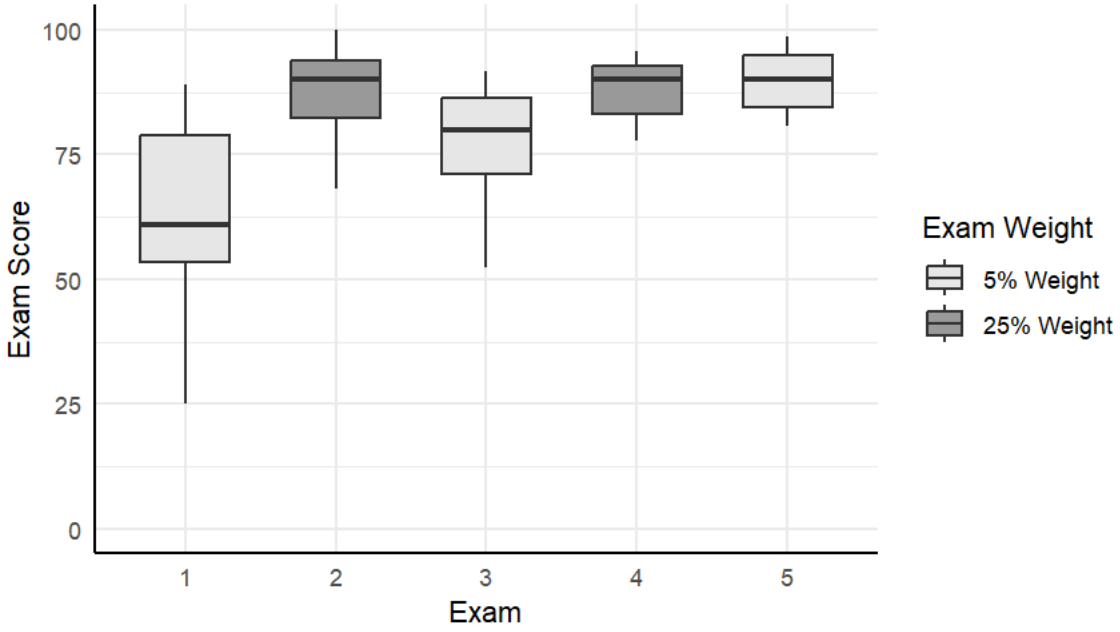
Predicted Learning and Well-Being Goal Ratings across Surveys in Experiment 4



Note. The learning goal ratings in Survey 12 was removed from the model due to external distractions (the presidential election). Shaded areas represent the 95% confidence intervals around the predicted values.

Figure 16

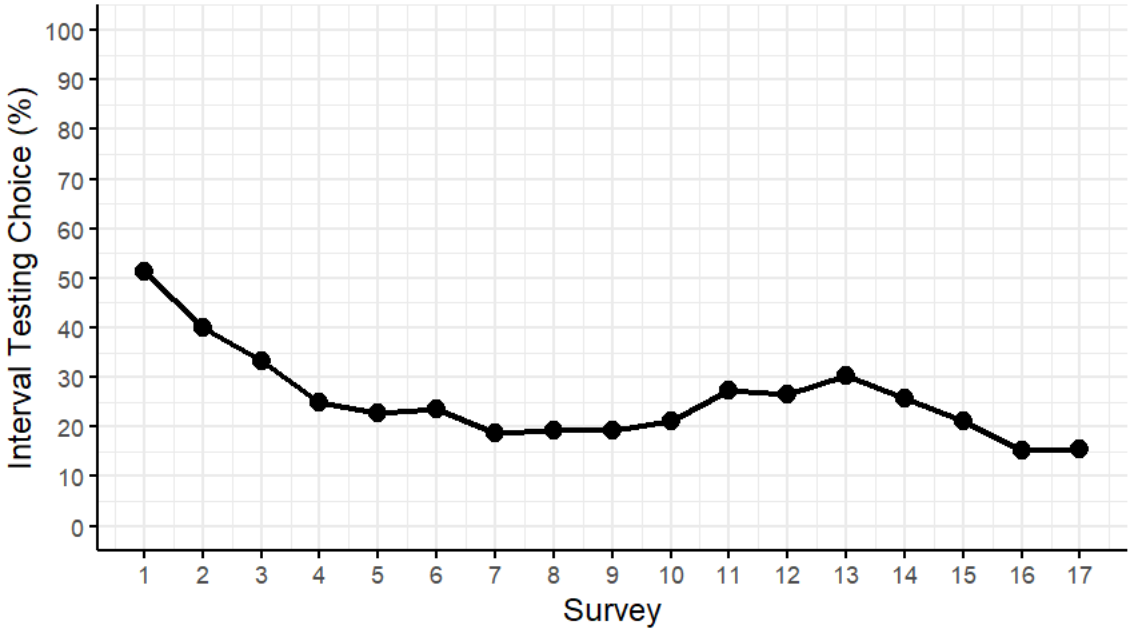
Exam Scores across Exams and Exam Weights in Experiment 4



Note. Exam 5 was a make-up exam, allowing participants the option to replace their Exam 1 score with their Exam 5 score if desired.

Figure 17

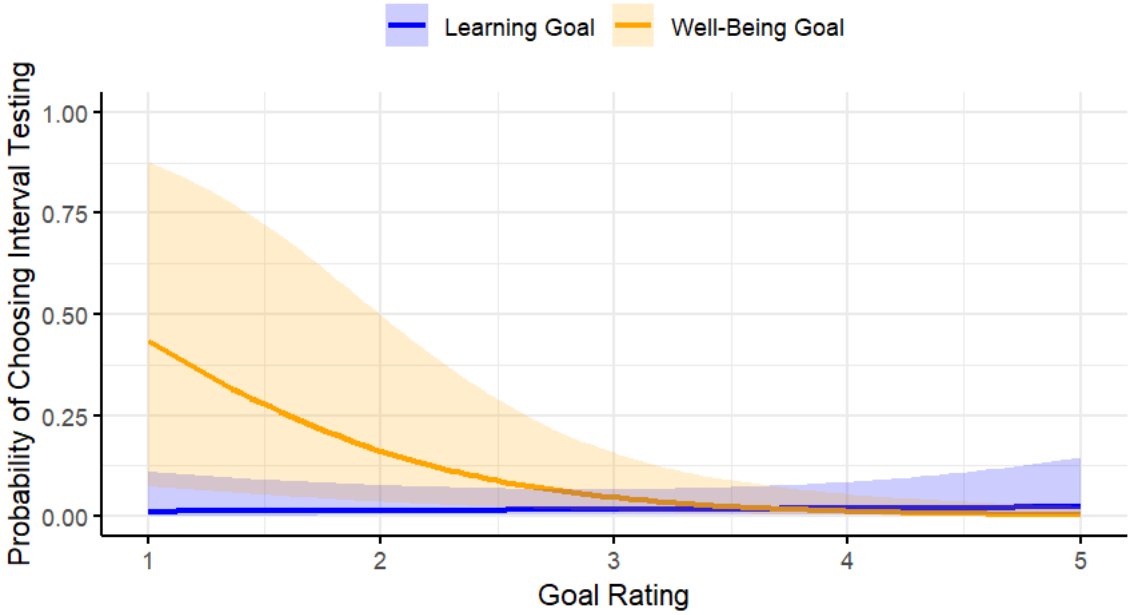
Percentage of Participants Choosing Interval Testing Across Surveys in Experiment 4



Note. Surveys 1 through 14 were conducted in person, while Surveys 15, 16, and 17 were completed remotely. The series of Memory Lectures began on the day participants completed Survey 10.

Figure 18

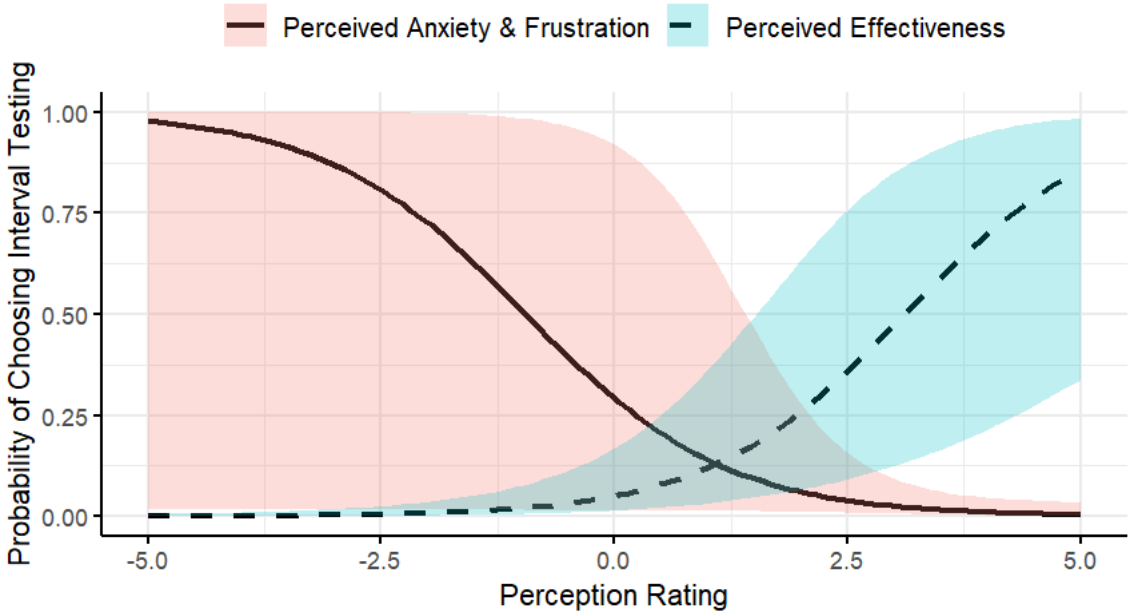
Predicted Choice of Practice Type as a Function of Learning and Well-Being Goals in Experiment 4



Note. Shaded areas represent the 95% confidence intervals around the predicted values.

Figure 19

Predicted Choice of Practice Type as a Function of Perceptions of Effectiveness, Anxiety and Frustration in Experiment 4



Note. Shaded areas represent the 95% confidence intervals around the predicted values.

Appendix A

Pre-Class Survey in Experiment 4

1. How do you typically study on your own? Please describe all of the study strategies you regularly use.

2. Self-testing Habits

	Almost Never	Sometimes	Often	Almost Always
How often do you test yourself when you study on your own?	1	2	3	4

3. Test Anxiety Inventory

Below is a list of statements that people have used to describe themselves. Please read each statement carefully and then select the number that best indicates how you generally feel. Remember, there are no wrong or right answers. Do not spend too much time on one statement but give the answer which seems to describe how you generally feel. Please answer every statement.

	Almost Never	Sometimes	Often	Almost Always
1. I feel confident and relaxed while taking tests.	1	2	3	4
2. While taking examinations, I have an uneasy, upset feeling.	1	2	3	4
3. Thinking about my grade in a course interferes with my work on tests.	1	2	3	4
4. I freeze up on important exams.	1	2	3	4

5. During exams, I find myself thinking about whether I'll ever get through school.	1	2	3	4
6. The harder I work at taking a test, the more confused I get.	1	2	3	4
7. Thoughts of doing poorly interfere with my concentration on tests.	1	2	3	4
8. I feel very jittery when taking an important test.	1	2	3	4
9. Even when I'm well prepared for a test, I feel very nervous about it.	1	2	3	4
10. I start feeling very uneasy just before getting a test paper back.	1	2	3	4
11. During tests, I feel very tense.	1	2	3	4
12. I wish examinations did not bother me so much.	1	2	3	4
13. During important tests, I am so tense that my stomach gets upset.	1	2	3	4
14. I seem to defeat myself while working on important tests.	1	2	3	4
15. I feel very panicky when I take an important test.	1	2	3	4
16. I worry a great deal before taking an important examination.	1	2	3	4
17. During tests, I find myself thinking about the consequences of failing.	1	2	3	4
18. I feel my heart beating very fast during important tests.	1	2	3	4
19. After an exam is over, I try to stop worrying about it, but I can't.	1	2	3	4
20. During examinations, I get so nervous that I forget facts I really know.	1	2	3	4

Appendix B

Mid-Semester Survey in Experiment 4

1. How many hours per week, outside of regularly scheduled class meetings, do you spend on this class?

- Less than 1 hour
- 1 – 2 hours
- 3 – 4 hours
- 5 or more hours

2. What study strategies do you use outside of the class meetings to learn the course material?

3. During this course, you have had the option to participate in either interval study or interval testing. Which strategy did you choose more often?

4. Why did you choose [the selected strategy] more often?

5. Based on your experience with these activities, which study strategy—interval testing or interval study—do you think is more likely to induce anxiety and frustration while studying?

6. How much more likely do you think [the selected strategy] is to cause anxiety and frustration compared to [the unselected strategy]?

- 1 - Slightly more likely
- 2 - Somewhat more likely
- 3 - Moderately more likely
- 4 - Much more likely
- 5 - Extremely more likely

7. Based on your experience with these activities, which study strategy—interval testing or interval study—do you think is more effective for learning?

8. How much more effective do you think [the selected strategy] is compared to [the unselected strategy]?

- 1 - Slightly more effective
- 2 - Somewhat more effective
- 3 - Moderately more effective
- 4 - Much more effective
- 5 - Extremely more effective

Appendix C

Course Syllabus of Cognitive Psychology (Fall 2024)

Course Time: Monday & Wednesday, 4:30–5:45 PM

Instructor: Dr. Ayanna K. Thomas

Teaching Assistant: Lan Anh Do, Graduate Student

Required Text: Cognition by Chun & Most

COURSE OVERVIEW

This course provides an in-depth exploration of classic and current issues in human cognition, examining behavioral, neuropsychological, and neuroscientific approaches to data and theory. Emphasis will be placed on theoretical models grounded in empirical support. Topics will include experimental and neuropsychological methodologies in cognitive science, sensation, perception, attention, working memory, encoding and retrieval processes, implicit memory and multiple memory systems, categorization, decision making, developmental changes in cognition, and language acquisition and comprehension.

COURSE INFORMATION

Students should come to class FULLY PREPARED, having read the complete assignment in *advance* and ready to DISCUSS the readings in class.

Class periods are used to present lecture material and facilitate discussions related to the assigned readings. Completing readings and assignments on time, coming to class, and actively participating in the learning process are essential. Students are encouraged to take detailed notes of all readings and lectures/discussions. Attendance is a must and is factored into your final grade. Course schedules are often tight and require students to maintain disciplined attendance and preparation. If a student must miss a class, that student must arrange to find out what was

covered in class, including news or changes in scheduled items. The Teaching Assistant is an excellent source for this information.

This Course is an Experiment

The course is structured to examine the effects of repeated and interval study and testing on learning. During each class meeting, I will lecture for approximately 20-30 minutes. Immediately after lecture, you will either be given a study sheet to review (interval study) or a short test (interval testing). You will have the option to engage in either interval study or interval testing. Typically, there will be two interval practice sessions – either study or testing – during each class meeting. After a series of interval practice and lectures, you will be given a cumulative exam. The cumulative exam will cover information from the previous series of lectures. Cumulative exams are only cumulative within a given unit.

Exams are Important

EXAMS

Unit Cumulative Exams will consist of 20-30 questions.

All exams will be given during class time, between lecture segments. 10 minutes will be allotted for interval practice. 30-45 minutes will be allotted for unit cumulative exams.

Exams will consist of multiple choice and short answer questions.

ACTIVITY

An in-class activity and discussions are planned for days where there are also unit exams scheduled. There is one class activity that replaces a class meeting.

GRADES are based on class attendance, engagement and exam performance.

20 class meetings without exams: 20%

Unit Exam 1: 5%

Unit Exam 2: 25%

Unit Exam 3: 5%

Unit Exam 4: 25%

Unit Exam 5: 5% (Optional — may be taken as a make-up for Exam 1, if desired)

Classroom Activities (4 total): 20% (5% each)

MISSING CLASS MEETINGS:

Students who will have to miss a class due to some scheduled event (religious observance, team meet, etc.) are required to notify me at the beginning of the semester during the first week of classes. Students will be allowed up to two missed classes without penalty. Secondary assignments will be used in these cases.

There are no options for make-up exams. However, if there is an emergency and you must miss an exam, please meet with me so we can discuss options.

GUIDELINES FOR SUCCESS

As the students in this class, you also have responsibilities. In order for you to learn the material and achieve the grade you desire I expect that you understand the following:

GRADES ARE EARNED. Thus, you have to work for high grades.

Show up for all classes.

Carefully follow the syllabus.

CLASS SCHEDULE

Week 0

Wed, Sept 4 — Course Overview, Informed Consent (Chapter 1)

Week 1

Mon, Sept 9 — History (Chapters 1 & 2)

Wed, Sept 11 — Goals & Practice Choice; Neural Basis of Vision (Chapters 2 & 3)

Week 2

Mon, Sept 16 — Perception (Chapter 3)

Wed, Sept 18 — Cumulative Exam 1 (5%) + Activity 1

Week 3

Mon, Sept 23 — Goals & Practice Choice; Faces & Object Recognition (Chapter 3)

Wed, Sept 25 — Goals & Practice Choice; Mental Imagery & Attention (Chapters 3.3 & 4.1–4.2)

Week 4

Mon, Sept 30 — Goals & Practice Choice; Attention & Inhibition (Chapters 4.3–4.6)Wed, Oct 2 — Practice Choice; In-Class Review

Week 5

Mon, Oct 7 — Cumulative Exam 2 (25%) + Activity 2

Wed, Oct 9 — Goals & Practice Choice; Working Memory (Chapter 5.2–end)

Week 6

Mon, Oct 14 — No Class (Indigenous Peoples Day)

Wed, Oct 16 — Goals & Practice Choice; Knowledge (Chapter 11.1)

Week 7

Mon, Oct 21 — Goals & Practice Choice; Knowledge continued

Wed, Oct 23 — Goals & Practice Choice; In-Class Review

Week 8

Mon, Oct 28 — Cumulative Exam 3 (5%) + Activity 3

Wed, Oct 30 — Goals & Practice Choice; Everyday Memory (Chapter 6)

Week 9

Mon, Nov 4 — Goals & Practice Choice; Everyday Memory (Chapter 6)

Wed, Nov 6 — Goals & Practice Choice; Complex Memories (Chapters 6 & 7)

Week 10

Mon, Nov 11 — No Class (Veterans Day)

Tues, Nov 12 — Goals & Practice Choice; Memory Systems (Chapter 7)

Wed, Nov 13 — Goals & Practice Choice; In-Class Review

Week 11

Mon, Nov 18 — Cumulative Exam 4 (25%) + Activity 4

Wed, Nov 20 — Goals & Practice Choice; Judgment & Decision Making (Chapter 9)

Week 12

Mon, Nov 25 — Goals & Practice Choice; Judgment & Decision Making continued

Wed, Nov 27 — No Class (Thanksgiving Break)

Week 13

Mon, Dec 2 — Goals & Practice Choice; Reasoning (Chapter 10)

Wed, Dec 4 — Course Review & Experiment Explanation

Week 14

Mon, Dec 9 — Cumulative Exam 5 (5%)

References

Agarwal, P. K., Karpicke, J. D., Kang, S. H. K., Roediger, H. L., & McDermott, K. B. (2008). Examining the testing effect with open- and closed-book tests. *Applied Cognitive Psychology*, *22*, 861-876.

Ames, C., & Archer, J. (1988). Achievement goals in the classroom: Students' learning strategies and motivation processes. *Journal of Educational Psychology*, *80*, 260-267.

Ariel, R., & Karpicke, J. D. (2018). Improving self-regulated learning with a retrieval practice intervention. *Journal of Experimental Psychology: Applied, 24*(1), 43–56. https

Ariel, R., Dunlosky, J., & Bailey, H. (2009). Agenda-based regulation of study-time allocation: When agendas override item-based monitoring. *Journal of Experimental Psychology: General, 138*(3), 432–447.

Badali, S., Rawson, K. A., & Dunlosky, J. (2022). Do students effectively regulate their use of self-testing as a function of item difficulty? *Educational Psychology Review, 34*, 1651–1677.

Bangert-Drowns, R. L., Kulik, C. L. C., Kulik, J. A., & Morgan, M. (1991). The instructional effect of feedback in test-like events. *Review of educational research*, *61*(2), 213-238.

Barnett, S. M., & Ceci, S. J. (2002). When and where do we apply what we learn?: A taxonomy for far transfer. *Psychological bulletin*, *128*(4), 612-637.

Bertsch, Sharon, Bryan J. Pesta, Richard Wiscott, and Michael A. McDaniel. "The generation effect: A meta-analytic review." *Memory & cognition* 35 (2007): 201-210.

Bjork, E. L., & Bjork, R. A. (2011). Making things hard on yourself, but in a good way: Creating desirable difficulties to enhance learning. *Psychology and the real world: Essays illustrating fundamental contributions to society, 2*(59-68).

Bjork, R. A. (1975). Retrieval as a memory modifier. In R. Solso (Ed.), Information processing and cognition: The Loyola Symposium, pp. 123–144. Hillsdale, NJ: Erlbaum.

Bjork, R. A., & Linn, M. C. (2006). The science of learning and the learning of science. *Aps Observer*, *19*(3).

Boekaerts, M. (2011). "Emotions, emotion regulation, and self-regulation of learning," in Handbook of Self-Regulation of Learning and Performance, eds B. J. Zimmerman and D. H. Schunk (New York, NY: Routledge), 408–425.

Boekaerts, M., & Cascallar, E. (2006). How far have we moved toward the integration of theory and practice in self-regulation?. *Educational psychology review*, *18*, 199-210.

Bottiroli, S., Dunlosky, J., Guerini, K., Cavallini, E., & Hertzog, C. (2010). Does task affordance moderate age-related deficits in strategy production?. *Aging, Neuropsychology, and Cognition*, *17*(5), 591-602.

Butler, A. C. (2010). Repeated testing produces superior transfer of learning relative to repeated studying. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *36*(5), 1118-1133.

Butler, A. C., & Roediger, H. L. (2008). Feedback enhances the positive effects and reduces the negative effects of multiple-choice testing. *Memory & cognition*, *36*(3), 604-616.

Carpenter, S. K. (2009). Cue strength as a moderator of the testing effect: the benefits of elaborative retrieval. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *35*(6), 1563-1569.

Carpenter, S. K., & DeLosh, E. L. (2006). Impoverished cue support enhances subsequent retention: Support for the elaborative retrieval explanation of the testing effect. *Memory & cognition*, *34*, 268-276.

Cassady, J. C. (2004). The influence of cognitive test anxiety across the learning–testing cycle. *Learning and instruction*, *14*(6), 569-592.

Castel, A. D., Murayama, K., Friedman, M. C., McGillivray, S., & Link, I. (2013). Selecting valuable information to remember: Age-related differences and similarities in self-regulated learning. *Psychology and Aging, 28,* 232–242.

Caviola, S., Carey, E., Mammarella, I. C., & Szucs, D. (2017). Stress, time pressure, strategy selection and math anxiety in mathematics: A review of the literature. *Frontiers in psychology*, *8*, 1488.

Cepeda, N. J., Coburn, N., Rohrer, D., Wixted, J. T., Mozer, M. C., & Pashler, H. (2009). Optimizing distributed practice: Theoretical analysis and practical implications. *Experimental psychology*, *56*(4), 236-246.

Chen, A., & Liu, X. (2009). Task values, cost, and choice decisions in college physical education. *Journal of Teaching in Physical Education*, *28*(2), 192-213.

Chun, M. M., & Most, S. B. (2021). *Cognition.* Oxford University Press.

Clark, D. A., & Svinicki, M. (2015). The effect of retrieval on post-task enjoyment of studying. *Educational Psychology Review*, *27*, 51-67.

Clark, D. A., & Svinicki, M. (2015). The effect of retrieval on post-task enjoyment of studying. *Educational Psychology Review*, *27*, 51-67.

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

Eccles J. S., Adler, T. F., Futterman, R., Goff, S. B., Kaczala, C. M., Meece, J. L., & Midgley, C. (1983). Expectancies, values, and academic behaviors. In J. T. Spence (Ed.), Achievement and achievement motivation (pp. 75–146).

Efklides, A. (2011). Interactions of metacognition with motivation and affect in self-regulated learning: The MASRL model. *Educational psychologist*, *46*(1), 6-25.

Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G* Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior research methods, 39*(2), 175-191.

Flake, J. K., Barron, K. E., Hulleman, C., McCoach, B. D., & Welsh, M. E. (2015). Measuring cost: The forgotten component of expectancy-value theory. *Contemporary educational psychology, 41*, 232-244.

Greene, B. A., Miller, R. B., Crowson, H. M., Duke, B. L., & Akey, K. L. (2004). Predicting high school students' cognitive engagement and achievement: Contributions of classroom perceptions and motivation. *Contemporary educational psychology*, *29*(4), 462-482.

Halamish, V., Madmon, I., & Moed, A. (2019). Motivation to learn: The long-term mnemonic benefit of curiosity in intentional learning. *Experimental Psychology, 66*(5), 319–330.

Harackiewicz, J. M., Barron, K. E., Carter, S. M., Lehto, A. T., & Elliot, A. J. (1997). Predictors and consequences of achievement goals in the college classroom: Maintaining interest and making the grade. *Journal of Personality and Social Psychology*, *73*, 1284-1295.

Hartwig, M. K., & Dunlosky, J. (2012). Study strategies of college students: Are self-testing and scheduling related to achievement?. *Psychonomic bulletin & review*, *19*, 126-134.

Higham, P. A., Zengel, B., Bartlett, L. K., & Hadwin, J. A. (2022). The benefits of successive relearning on multiple learning outcomes. *Journal of Educational Psychology*, *114*(5), 928-944.

Hinze, S. R., & Rapp, D. N. (2014). Retrieval (sometimes) enhances learning: Performance pressure reduces the benefits of retrieval practice. *Applied Cognitive Psychology*, *28*(4), 597-606.

Jacoby, L. L. (1978). On interpreting the effects of repetition: Solving a problem versus remembering a solution. *Journal of Verbal Learning & Verbal Behavior, 17,* 649-667.

Janes, J. L., Dunlosky, J., & Rawson, K. A. (2018). How do students use self-testing across multiple study sessions when preparing for a high-stakes exam?. *Journal of Applied Research in Memory and Cognition*, *7*(2), 230-240.

Kadioglu, C., & Kondakci, E. U. (2014). Relationship between learning strategies and goal orientations: A multilevel analysis. *Eurasian Journal of Educational Research*, *56*, 1-22.

Kang, S. H., & Pashler, H. (2012). Learning painting styles: Spacing is advantageous when it promotes discriminative contrast. *Applied Cognitive Psychology*, *26*(1), 97-103.

Karpicke, J. D., & Roediger III, H. L. (2007). Repeated retrieval during learning is the key to long-term retention. *Journal of memory and language, 57*(2), 151-162.

Karpicke, J. D., & Roediger III, H. L. (2008). The critical importance of retrieval for learning. *Science, 319*(5865), 966-968.

Kim, Y. E., Yu, S. L., Wolters, C. A., & Anderman, E. M. (2023). Self-regulatory processes within and between diverse goals: the multiple goals regulation framework. *Educational Psychologist*, *58*(2), 70-91.

Kirk-Johnson, A., Galla, B. M., & Fraundorf, S. H. (2019). Perceiving effort as poor learning: The misinterpreted-effort hypothesis of how experienced effort and perceived learning relate to study strategy choice. *Cognitive psychology*, *115*, 101237.

Koriat, A., & Bjork, R. A. (2006). Mending metacognitive illusions: a comparison of mnemonic-based and theory-based procedures. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *32*(5), 1133-1145.

Koriat, A., Ma'ayan, H., & Nussinson, R. (2006). The intricate relationships between monitoring and control in metacognition: Lessons for the cause-and-effect relation between subjective experience and behavior. *Journal of Experimental Psychology: General, 135*(1), 36–69.

Kornell, N. (2009). Optimising learning using flashcards: Spacing is more effective than cramming. *Applied Cognitive Psychology: The Official Journal of the Society for Applied Research in Memory and Cognition*, *23*(9), 1297-1317.

Kornell, N., & Bjork, R. A. (2007). The promise and perils of self-regulated study. *Psychonomic bulletin & review*, *14*(2), 219-224.

Kornell, N., & Bjork, R. A. (2008). Learning concepts and categories: Is spacing the "enemy of induction"?. *Psychological Science*, *19*(6), 585-592.

Kornell, N., & Son, L. K. (2009). Learners' choices and beliefs about self-testing. *Memory, 17*(5), 493-501.

Kuhbandner, C., & Emmerdinger, K. J. (2019). Do students really prefer repeated rereading over testing when studying textbooks? A reexamination. *Memory*, *27*(7), 952-961.

Kulhavy, R. W., & Stock, W. A. (1989). Feedback in written instruction: The place of response certitude. *Educational psychology review*, *1*, 279-308.

Liu, S., Zhao, W., Shanks, D. R., Hu, X., Luo, L., & Yang, C. (2024). Effects of test anxiety on self-testing and learning performance. *Educational Psychology Review*, *36*(2), 59.

Lyle, K. B., & Crawford, N. A. (2011). Retrieving essential material at the end of lectures improves performance on statistics exams. *Teaching of Psychology*, *38*(2), 94-97.

Mawhinney, V. T., Bostow, D. E. , Laws, D. R., Blumenfeld, G. J., & Hopkins, B. L. (1971). A comparison of students studying-behavior produced by daily, weekly, and three week testing schedules. *Journal of Applied Behavior Analysis, 4,* 257–264.

McCabe, J. (2011). Metacognitive awareness of learning strategies in undergraduates. *Memory & cognition*, *39*, 462-476.

McDaniel, M. A., & Einstein, G. O. (2020). Training learning strategies to promote self-regulation and transfer: The knowledge, belief, commitment, and planning framework. *Perspectives on Psychological Science, 15*, 1363–1381.

McDaniel, M. A., Anderson, J. L., Derbish, M. H., & Morrisette, N. (2007). Testing the testing effect in the classroom. *European journal of cognitive psychology*, *19*(4-5), 494-513.

McDaniel, M. A., Fadler, C. L., & Pashler, H. (2013). Effects of spaced versus massed training in function learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *39*(5), 1417-1432.

Michael, J. (1991). A behavioral perspective on college teaching. *The Behavioral Ana lyst, 14,* 229–239.

Miller, T. M., & Geraci, L. (2014). Improving metacognitive accuracy: How failing to retrieve practice items reduces overconfidence. *Consciousness and cognition*, *29*, 131-140.

Morehead, K., Rhodes, M. G., & DeLozier, S. (2016). Instructor and student knowledge of study strategies. *Memory*, *24*(2), 257-271.

Muenks, K., Wigfield, A., Yang, J. S., & O'Neal, C. R. (2017). How true is grit? Assessing its relations to high school and college students' personality characteristics, self-regulation, engagement, and achievement. *Journal of Educational Psychology, 109*(5), 599–620.

Orfus, S. (2008). *The effect (of) test anxiety and time pressure on performance. The Huron University College Journal of Learning and Motivation, 46 (1), 7*.

Ott, M. K. (2017). *Testangst und deren Zusammenhang mit Testleistung: Effekte von Messzeitpunkt, Instruktion und funktionaler Bewertung der Angst [Test anxiety and its relationship to test performance: effects of time of measurement, instruction and functional assessment of anxiety]* (Doctoral dissertation, Justus-Liebig-Universität Giessen).

Panadero, E. (2017). A review of self-regulated learning: Six models and four directions for research. *Frontiers in psychology*, *8*, 422.

Pashler, H., Cepeda, N. J., Wixted, J. T., & Rohrer, D. (2005). When does feedback facilitate learning of words?. *Journal of experimental psychology: Learning, Memory, and Cognition*, *31*(1), 3-8.

Phan, H. P. (2009). Amalgamation of future time orientation, epistemological beliefs, achievement goals and study strategies: Empirical evidence established. *British Journal of Educational Psychology*, *79*, 155-173.

Pintrich, P. R., & DeGroot, E. V. (1990). Motivational and self-regulated learning components of classroom academic performance. *Journal of Educational Psychology, 82,* 33-40.

Pintrich, P. R., & Schunk, D. H. (2002). *Motivation in education: Theory, research, and applications*. Upper Saddle River, NJ: Prentice Hall.

Printablee. (n.d.). *Adult hidden object worksheets*. Printablee.com. https://www.printablee.com/post_adult-hidden-object-printable-worksheets_258545/

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

Rea, S. D., Wang, L., Muenks, K., & Yan, V. X. (2023). Students can (mostly) recognize effective learning, so why do they not do it?. *Journal of Intelligence*, *10*(4), 127.

Rivers, M. L. (2021). Metacognition about practice testing: A review of learners' beliefs, monitoring, and control of test-enhanced learning. *Educational Psychology Review*, *33*(3), 823-862.

Rivers, M. L. (2021). Metacognition about practice testing: A review of learners' beliefs, monitoring, and control of test-enhanced learning. *Educational Psychology Review, 33*(3), 823–862.

Roediger III, H. L., & Karpicke, J. D. (2006b). The power of testing memory: Basic research and implications for educational practice. *Perspectives on psychological science*, *1*(3), 181-210.

Roediger III, H. L., Putnam, A. L., & Smith, M. A. (2011). Ten benefits of testing and their applications to educational practice. *Psychology of learning and motivation*, *55*, 1-36.

Roediger, H. L., & Karpicke, J. D. (2006a). Test enhanced learning: Taking memory tests improves long-term retention. *Psychological Science*, *17*, 249-255.

Rowland, C. A. (2014). The effect of testing versus restudy on retention: a meta-analytic review of the testing effect. *Psychological bulletin, 140*(6), 1432-1463.

Sarason, I. G., & Sarason, B. R. (1990). Test anxiety. In *Handbook of Social and Evaluation Anxiety* (pp. 475-495). Springer.

Simone, P. M., Whitfield, L. C., Bell, M. C., Kher, P., & Tamashiro, T. (2023). Shifting students toward testing: Impact of instruction and context on self-regulated learning. *Cognitive Research: Principles and Implications*, *8*(1), 14.

Slamecka, N. J., & Graf, P. (1978). The generation effect: Delineation of a phenomenon. *Journal of Experimental Psychology: Human Learning & Memory, 4,* 592-604.

Son, L. K. (2005). Metacognitive control: Children's short-term versus long-term study strategies. *Journal of General Psychology*, *132*, 347-363.

Spielberger, C. D. (1980). *Test anxiety inventory: Preliminary professional manual*. Consulting Psychologists Press. Spielberger, C. D.,

Thomas, A. K., Smith, A. M., Kamal, K., & Gordon, L. T. (2020). Should you use frequent quizzing in your college course? Giving up 20 minutes of lecture time may pay off. *Journal of Applied Research in Memory and Cognition*, *9*(1), 83-95.

Thomas, A. K., Wulff, A. N., Landinez, D., & Bulevich, J. B. (2022). Thinking about thinking about thinking… & feeling: A model for metacognitive and meta-affective processes in task engagement. *Wiley Interdisciplinary Reviews: Cognitive Science*, *13*(6), e1618.

Toppino, T. C., LaVan, M. H., & Iaconelli, R. T. (2018). Metacognitive control in self-regulated learning: Conditions affecting the choice of restudying versus retrieval practice. *Memory & Cognition*, *46*, 1164-1177.

Tullis, J. G., & Maddox, G. B. (2020). *Self-reported use of retrieval practice varies across age and domain. Metacognition and Learning, 15 (2), 129–154*.

Vaughn, K. E., & Kornell, N. (2019). How to activate students' natural desire to test themselves. *Cognitive research: principles and implications, 4,* 1-16.

Wang, L., Muenks, K., & Yan, V. X. (2023). Interventions to promote retrieval practice: Strategy knowledge predicts intent, but perceived cost predicts usage. *Journal of educational psychology*, *115*(8), 1070-1086.

Wang, L., Muenks, K., & Yan, V. X. (2023). Interventions to promote retrieval practice: Strategy knowledge predicts intent, but perceived cost predicts usage. *Journal of Educational Psychology, 115*(8), 1070–1086.

Weissgerber, S. C., Reinhard, M.-A., & Schindler, S. (2016). Study harder? The relationship of achievement goals to attitudes and self-reported use of desirable difficulties in self-regulated learning. *Journal of Psychological and Educational Research*, *24*(1), 42–60.

Weissgerber, S. C., Reinhard, M.-A., & Schindler, S. (2018). Learning the hard way: Need for cognition influences attitudes towards and self-reported use of desirable learning difficulties. *Educational Psychology: An International Journal of Experimental Educational Psychology*, *38*(2), 176–202.

Wenzel, K. (2022). *Taking a Closer Look at the Desirability of Desirable Difficulties-Additionally Focusing on Required Prerequisites, Negative Side-Effects, and Negative Consequences* (Doctoral dissertation).

Wenzel, K., & Reinhard, M.-A. (2021a). Does the end justify the means? Learning tests lead to more negative evaluations and to more stress experiences. *Learning and Motivation, 73*, Article 101706.

Wenzel, K., & Reinhard, M.-A. (2021b). Learning with a double-edged sword? Beneficial and detrimental effects of learning tests—taking a first look at linkages among tests, later

learning outcomes, stress perceptions, and intelligence. *Frontiers in Psychology*, *12*, Article 693585.

Yan, V. X., & Schuetze, B. A. (2022). Not just stimuli structure: Sequencing effects in category learning vary by task demands. *Journal of Applied Research in Memory and Cognition*, *11*(2), 218-228.

Yeager, D. S., Romero, C., Paunesku, D., Hulleman, C. S., Schneider, B., Hinojosa, C., Lee, H. Y., O'Brien, J., Flint, K., Roberts, A., Trott, J., Greene, D., Walton, G. M., & Dweck, C. S. (2016). Using design thinking to improve psychological interventions: The case of the growth mindset during the transition to high school. *Journal of Educational Psychology, 108*, 374–391.