Educational Psychology Review
https://doi.org/10.1007/s10648-020-09562-w

REVIEW ARTICLE

Students' Use of Retrieval in Self-Regulated Learning: Implications for Monitoring and Regulating Effortful Learning Experiences

Shana K. Carpenter[1] **· Tino Endres**[2] **· Luotong Hui**[3]

Published online: 05 August 2020

Abstract
Retrieval practice has been widely studied as an effective strategy for enhancing memory. In this review article, we discuss how its effects on learning complex problem-solving procedures are less straightforward, however, with repeated studying of worked examples sometimes more effective than problem-solving practice. This worked example benefit has been interpreted within the framework of cognitive load theory. In both memory-based tasks and problem-solving tasks, students rate retrieval as more effortful, and less effective for learning, than repeated study. Self-regulated learning decisions do not align with the evidence about the effectiveness of retrieval, as students often avoid using retrieval in memory-based tasks but frequently use it in more complex problem-solving tasks. Patterns associated with self-report survey data and self-regulated learning decisions suggest that retrieval may be used primarily as a means of checking knowledge, and the effort experienced during retrieval may drive subsequent study decisions (i.e., the choice to engage in retrieval vs. further study) to the extent that the experienced effort is interpreted as a sign that learning has been ineffective. We discuss implications of students' views about the purpose of retrieval for effective monitoring and regulation of effort during learning, and propose interventions that may improve students' optimal use of retrieval in their study decisions.

Keywords Retrieval practice · Self-regulated learning · Cognitive load · Effort · Metacognition

Learners of today have a wide variety of options for how to go about acquiring new knowledge and skills. Non-traditional, unstructured learning opportunities such as massive open online courses (MOOCs) are becoming increasingly available, and formal educational environments as well are providing increased amounts of auxiliary materials and extra resources—such as

This material is based upon work supported by the James S. McDonnell Foundation 21st Century Science Initiative in Understanding Human Cognition, Collaborative Grant No. 220020483.

✉ Shana K. Carpenter
shacarp@iastate.edu

Extended author information available on the last page of the article

supplemental instruction, chapter quizzes, and online discussions—to help students learn. The ever increasing availability of options means that students have more choices now than ever before with regard to how they go about learning information.

Among the many choices that students face are what kind of learning resources and strategies to use, how to use them, and when. For example, students wishing to learn and remember information for exams might know that it is good to study the information more than once, but how many times should they study? When is the best time to study? What is the best way? Assuming the goal is to learn the information well, students would want to know that they are making the right study decisions that will maximize their understanding of the material and meet their learning goals.

Meeting these goals requires skill at successfully monitoring and regulating one's own learning. Research on self-regulated learning examines the decisions that students make while studying, what informs those decisions, and the consequences of those decisions for learning outcomes (Bjork et al. 2013; de Bruin and van Gog 2012; Dunlosky and Ariel 2011). One area of particular interest concerns the degree to which students are making the right decisions during study that will optimize their learning. That is, how well do students' study decisions align with strategies that are known to effectively enhance learning?

Several decades of research from cognitive science and educational psychology have revealed principles that reliably enhance learning (for reviews, see Dunlosky et al. 2013; Pashler et al. 2007; Rohrer and Pashler 2010). For example, learning can be significantly enhanced by inserting extra time in-between repeated study sessions—i.e., the widely documented benefits of *spacing*, or *distributed practice* (Carpenter 2017; Delaney et al. 2010; Toppino and Gerbier 2014)—by mixing up, or *interleaving*, the order of different types of procedures when studying things like mathematical formulas (Rohrer and Taylor 2007; Taylor and Rohrer 2010), categorical properties (Kornell and Bjork 2008), and language skills (Pan et al. 2019), and by practicing to retrieve information from memory rather than studying it over again (Rowland 2014). These techniques increase the difficulty or effort involved in learning while also increasing the ultimate durability of that learning, and as such they have come to be known as "desirable difficulties" (Bjork et al. 2013).

Students do not have a good sense of the effectiveness of these techniques, however. When asked to judge the effectiveness of these techniques for their own learning, students often express that they learn better through massed rather than distributed practice (Logan et al. 2012), through blocked rather than interleaved presentation (Kornell and Bjork 2008; Wahlheim et al. 2012), and through repeated studying rather than retrieval (Tullis et al. 2013). Even after direct experience—for example, by learning some material through blocking and some through interleaving—students still tend to endorse the less effective strategy of blocking, even when their own performance reveals benefits of interleaving (Birnbaum et al. 2013).

Thus, students have a hard time determining when effort is good for learning. As successful learning depends on investing the right amount of effort at the right times, the accurate monitoring and regulation of effort are critical to the learning process. Chief among the challenges involved in accomplishing this is the fact that effort is not always a diagnostic cue of learning. Although the effort involved in learning can sometimes reflect desirable difficulties, other times effort can arise from factors that do not benefit—and could significantly *impair*—learning. This places particular importance on understanding students' awareness of the distinction between effortful experiences that are productive vs. non-productive for learning, and students' regulation of cognitive resources in a way that optimally manages this effort to achieve learning goals.

The Effort Monitoring and Regulation (EMR) Framework integrates cognitive load theory and self-regulated learning theory to address these important issues (de Bruin et al. 2020). The framework incorporates input from the task level and metacognitive level that interact via monitoring and control processes, with cognitive load incorporated as a central component that interacts with both levels and both processes. Three broad research questions generated from this framework include (1) How do students monitor effort? (2) How do students regulate effort? and (3) How do we optimize cognitive load on self-regulated learning tasks?

In this article, we focus on Questions #1 and #2 within the context of a learning technique that has been widely studied but appears to be sub-optimally utilized in students' self-regulated learning decisions. *Retrieval practice*—the act of practicing to retrieve information from memory—can significantly enhance memory retention of that information. Although the benefits of retrieval practice are well-documented for memory-based tasks, these benefits are less likely to occur for more complex problem-solving tasks. Cognitive load theory provides a framework for predicting when the efforts involved in retrieval are beneficial and when they are not. Interestingly, students' self-regulated learning decisions involving retrieval are misaligned with the empirical evidence, as students tend to avoid using retrieval when it would be beneficial (for memory-based tasks) and they readily use retrieval when it would not always be beneficial (for problem-solving tasks). In both types of tasks students rate retrieval as more effortful, and less effective for learning, compared to pure study. We review evidence showing that students' self-regulation decisions involving retrieval may be driven by students' views about the role of retrieval in learning, and how they might interpret the role of effort involved in retrieval. Finally, relevant to Question #3 of the EMR framework, we propose potential intervention approaches that may increase students' likelihood of optimally using retrieval practice when they study on their own.

Retrieval Practice

Retrieval practice is one of the most widely studied cognitive science principles for enhancing memory. The earliest studies on retrieval practice were conducted over 100 years ago (e.g., Abott 1909; Gates 1917), and in the last few decades alone, hundreds of additional studies have been added to the ever growing literature on this topic (for recent reviews and meta-analyses, see Adesope et al. 2017; Dunlosky et al. 2013; Rowland 2014).

The typical design of a retrieval practice study involves an initial opportunity to encode some information (e.g., reading a text passage), followed by an opportunity to retrieve that information from memory. This retrieval condition can be compared to a control condition that involves additional exposure to the information without requiring retrieval, such as another opportunity to simply read the information. On a final memory test over all of the information given some time later, information learned through retrieval is typically remembered significantly better than information learned through re-reading.

This finding has been observed in numerous laboratory studies using a variety of materials such as simple word lists and word pairs (Carpenter et al. 2006; Carpenter and Yeung 2017), face-name pairs (Carpenter and DeLosh 2005), foreign language vocabulary (Pyc and Rawson 2010), trivia facts (Carpenter et al. 2008), text passages (Butler 2010; Hinze et al. 2013), and video-based materials (Butler and Roediger III 2007; Endres et al. 2017). The benefits of retrieval have been shown to be quite durable, as final tests that are given weeks or months later still show memory advantages of information that had been retrieved compared to restudied (Carpenter et al. 2009; Carpenter et al. 2008).

Additional research has revealed factors that sometimes moderate the benefits of retrieval. Although one retrieval (vs. restudy) opportunity can produce these benefits, the effects of retrieval are stronger when multiple retrieval (vs. restudy) opportunities are provided (Carpenter et al. 2008; Kuo and Hirshman 1996; Rawson and Dunlosky 2011). The benefits of retrieval over restudy can also occur even when no corrective feedback or opportunity to restudy the material after retrieval is provided (Carpenter 2009, 2011); however, retrieval is even more effective when followed by feedback that allows learners to correct errors (Kang et al. 2007). Part of what makes retrieval effective, therefore, could be that it enhances the effectiveness of additional encoding opportunities through revealing to learners what they know (i.e., can retrieve) and do not know (i.e., cannot retrieve). Indeed, several studies have found that metacognitive calibration—i.e., the accuracy with which learners can predict future test performance—is significantly enhanced following retrieval attempts compared to restudy (Agarwal et al. 2008; Carpenter et al. 2016; Little and McDaniel 2014; Tullis et al. 2013). Retrieval may thus have direct benefits on memory through the processes involved in the act of retrieval itself, as well as indirect benefits through the act of revealing to learners what they know and do not know, that enhance the effectiveness of subsequent encoding opportunities (Arnold and McDermott 2013; Roediger III et al. 2011b).

Studies conducted in real educational settings have confirmed the benefits of retrieval. For example, retrieval (relative to re-reading or re-exposure to the information) has been shown to significantly benefit memory for the names of toys in preschool children (Fritz et al. 2007), new vocabulary in elementary school students (Goossens et al. 2014a; Goossens et al. 2014b), science information in third grade students (Jaeger et al. 2015), social studies information in middle school students (Roediger III et al. 2011a), history information in middle school and high school students (Carpenter et al. 2009; McDermott et al. 2014), and psychology information in college students (McDaniel et al. 2007; Thomas et al. 2020). Real classroom environments are not always conducive to the inclusion of a re-exposure control group, so some classroom-based studies on retrieval practice have instead been designed to compare the effects of retrieval activities vs. no retrieval activities (e.g., McDaniel et al. 2011). Studies of both types have been accumulated in recent reviews and meta-analyses reporting beneficial effects of retrieval in classroom-based settings (Adesope et al. 2017; Schwieren et al. 2017).

Limitations of Retrieval Practice

Studies on retrieval practice have typically explored memory retention of relatively discrete types of material that can be straightforwardly recalled. Though the benefits of retrieval are well-established for materials such as word pairs, foreign language vocabulary, definitions, and recall of text materials, the effects of retrieval have been explored less often with more complex tasks that require learners to not just recall information, but also apply it in a new way.

Some studies have begun to explore the effects of retrieval on the formation and application of complex concepts and principles, and have found that the effects of retrieval in these types of tasks are not always straightforward. Tran et al. (2015) explored the effects of retrieval vs. rereading on deductive inferences. They first presented participants with scenarios that consisted of several premises (e.g., (1) students commute from off-campus housing to campus by any of 3 routes; (2) on Ash, students may walk, bike, or drive; (3) on Birch, only walking is allowed; (4) on Canyon, only biking or walking is allowed; (5) if it is raining, students do not bike; (6) if it is hot, students drive to campus or take Birch; (7) freshmen are not allowed to have cars). After reading the premises once, participants either continued reading the premises again (rereading), or they tried to recall portions of the premises that were missing (retrieval; e.g., On Birch, only _______ is allowed) and then received feedback of the correct answer (e.g., walking). On both immediate and delayed final tests requiring participants to draw inferences from the premises (e.g., Which form(s) of transportation are allowed on all three routes? (A) walking, (B) biking, (C) driving, (D) walking and driving, (E) walking and biking), there was no difference in performance according to whether participants learned the premises through rereading or through retrieval.

Follow-up studies have revealed that retrieval can benefit deductive reasoning when it promotes more effective memory for the premises, however. Wissman et al. (2018) observed the same outcomes as Tran et al. (2015) when using the same materials and same basic procedure. However, when participants were given extra practice at retrieving the premises through cued recall with feedback, initial retrieval success was increased, and subsequent final test performance on both a memory test and deductive inference test was higher following retrieval compared to rereading. Eglington and Kang (2018) also found the same results as Tran et al. when the premises were presented one at a time. However, when the premises were presented simultaneously (to facilitate processing of the relationships between them), retrieval facilitated later performance on the deductive inference test.

Other studies have explored the effects of retrieval on analogical problem-solving. Peterson and Wissman (2018) provided participants with a problem scenario about a military general who plans to lead his army to capture the fortress of an evil dictator, but must avoid the landmines that the dictator has placed along the roads leading to the fortress (see Gick and Holyoak 1980). The mines can be avoided if small numbers of people traverse the roads, but not if the entire army traverses at once. The solution to the problem (divide the army into smaller groups that each uses a different road) was presented along with the problem scenario. After reading the problem and solution, participants either read it again or attempted to retrieve it from memory. When given a new problem to solve that differed in surface features but required the same underlying solution (e.g., a doctor wishing to use radiation to destroy a tumor, but the intensity needed to destroy the tumor would also destroy healthy tissue), participants were not generally better at solving the new problem if they had previously retrieved (rather than reread) the original problem and solution. However, when participants performed well on the initial retrieval task (i.e., when they retrieved enough key idea units needed to solve the problem), those who retrieved rather than reread the initial problem and solution were significantly more likely to solve the second problem.

Using similar materials, Hostetter et al. (2019) found that retrieval of a problem scenario (compared to rereading or copying the scenario) did not enhance later ability to apply the same solution to a new problem. In their design, the final test included new problems to solve in addition to a memory test over the original problem scenario that was learned via retrieval, rereading, or copying. Although retrieval practice benefited overall memory for the original problem scenario, it did not enhance memory for the specific details that were common to the original and new problem and that were critical for solving the new problem (Experiment 1). However, when participants learned the original scenario by retrieving or rereading only the details that were critical for solving the new problem they would later receive (Experiment 2), memory for these details was enhanced by retrieval, as was the ability to solve the new problem.

Thus, as with deductive inference tasks, the potential of retrieval to enhance analogical problem-solving appears to depend partly upon the degree to which it enhances memory for the requisite information needed to solve the problem. Both Peterson and Wissman (2018) and

Hostetter et al. (2019) found that retrieval of the original problem only benefited solving of the new problem when participants were given a hint that the previously-learned solution could be used to solve the new problem. This indicates that although retrieval may enhance memory for the information needed to solve a problem, it does not appear to enhance the spontaneous decision to use that information in a new context.

These findings suggest that retrieval may enhance performance on complex tasks to the extent that performance depends, at least in part, on memory of previously acquired information. Although this seems to be the case for deductive and analogical reasoning, other types of complex tasks may be less dependent upon memory for specific items. The learning of concepts and principles, for example, requires learners to build a general representation of those concepts or principles through identifying their common properties, which might rely on the extraction of commonalities across multiple instances rather than on memory for specific pieces of information. To the extent that retrieval practice primarily benefits memory retention, retrieval may not be the optimal way to facilitate the processing needed for the induction of complex concepts and principles.

Studies on the acquisition of problem-solving procedures have explored the ways in which learning is affected by studying examples of the procedures vs. trying to use those procedures to solve problems. To the extent that problem-solving attempts represent a form of retrieval (i.e., trying to apply a previously-introduced solution to a current problem), this research shows that those retrieval attempts are generally not more effective than opportunities for extended study of the solution procedures. van Gog et al. (2011) presented a series of problem scenarios on electrical circuit malfunctions. Students' job was to learn to diagnose and troubleshoot the malfunctions. During a learning phase, the scenarios were sometimes presented in the form of problems that needed to be solved. That is, students were presented with a diagram of an electric circuit and had to apply a particular law to determine how the circuit should function. Then, based on a series of measurements from the system that indicated it was not functioning properly, they had to diagnose where the malfunction was occurring. Other times, the same scenarios were presented as fully worked out examples with all of the information provided about how the circuit should function and where the malfunction occurred, so that students could study the problem and solution procedures but did not have to try to solve it. Students were randomly assigned to the problem solving condition or to the worked example condition, or two hybrid conditions in which problem solving was followed by a worked example (problem-example) or a worked example was followed by an attempt to solve the problem (example-problem). On an immediate final test over new problems that had the same type of circuit malfunctions, performance was best for students who learned from worked examples compared to problem solving. Worked examples followed by problem solving (example-problem) produced performance comparable to worked examples alone, and problem solving followed by worked examples (problem-example) produced performance comparable to problem solving. Furthermore, ratings of mental effort during learning were higher for the problem solving and problem-example groups than for the worked example and example-problem groups, indicating that starting with problem-solving practice (rather than worked examples) was more effortful.

Using the same materials, van Gog and Kester (2012) found that immediate final test performance was similar following study of fully worked examples vs. example-problem sequences. van Gog et al. (2015) replicated this same outcome on both an immediate and 1-week delayed final test, as well as the finding reported by van Gog et al. (2011) that mental effort during learning was higher for example-problem sequences than for worked examples.

Similar results have been found for different types of problem solving tasks. Baars et al. (2014) found no significant differences in the degree to which fifth grade students learned to solve the “water jug” problem as a function of whether they learned from worked examples or alternating example-problem pairs. This was the case even though the latter provided a greater number of exposures to the material and thus more opportunities for practice. Judgments of learning also showed that students overestimated their own future test performance to a greater degree after worked examples compared to problem-solving. Using probability problems based on Poisson distributions, Yeo and Fazio (2019, Experiment 1) found that studying four worked examples, compared to one worked example followed by three problem-solving attempts, produced significantly better problem-solving performance on final tests administered after both 5 min and 1 week. Judgments of learning were also higher following study of worked examples compared to problem-solving practice. The benefits of worked examples have been observed in a variety of other problem-solving tasks as well (e.g., Cooper and Sweller 1987; for reviews, see Atkinson et al. 2000; Renkl 2014; Salden et al. 2010; Sweller 2006).

Cognitive Load Theory

The pattern of results described above has been interpreted within the framework of cognitive load theory (Paas et al. 2003; Paas et al. 2010; Sweller 1988, 1994). This framework proposes that learners have limited cognitive resources, and the way in which those resources are expended during learning determines whether or not learning will be effective. Learning is optimized to the extent that these limited resources are spent processing the relevant aspects of the task that are needed for long-term learning.

Processing resources are affected by cognitive load, which can arise from intrinsic sources (inherent to the type of learning or the individual learner’s knowledge), extraneous sources (factors present in a learning situation that do not contribute to learning), or germane sources (relevant aspects of the task that contribute directly to learning) (Sweller et al. 1998; Sweller 2010). Learning suffers if working memory is overloaded by processing information that is not relevant to the knowledge and skills required for the task. Complex learning situations contain a variety of elements that must be processed, and successful learning depends upon whether these elements are processed in a way that optimally manages cognitive resources, minimizing extraneous load and allowing sufficient room for germane load. The complexity of a learning situation is defined by the level of element interactivity, or the number of elements that must be simultaneously processed and related to one another in working memory (Sweller 1994, 2010). Some types of learning are inherently low in element interactivity, such as the learning of individual words or vocabulary that can occur without reference to other individual items being learned. Other types of learning, such as a series of interrelated steps to carrying out a solution to a problem, are higher in element interactivity because they contain multiple pieces of information that must be integrated and processed simultaneously.

Cognitive load theory is commonly used to explain the effects of worked examples in learning complex concepts and principles. In the early stages of learning, novice learners benefit by viewing worked examples because these examples illustrate the principle being learned and encourage limited resources to be spent on processing the relevant properties of that principle that will foster the development of a general representation or schema (Kalyuga et al. 2010; Sweller 1988, 2010). Attempting to apply the principle too early (before learners have developed an understanding of it) encourages the use of strategies that make ineffective use of limited resources (e.g., trial-and-error or means-end analysis), divert attention away from developing an understanding of the principle, and undermine schema acquisition. Once learners develop a schema for the principle, they can shift resources to processing the relevant principle-based information and no longer expend resources on irrelevant information. Worked examples that illustrate that principle now become redundant and resources are more effectively spent on practicing to apply the principle. Consistent with this reasoning, studies have shown that the effectiveness of worked examples depends on a learner's level of knowledge. Novice learners who have low prior knowledge on a given topic learn as well or better by starting with worked examples compared to problem-solving practice, whereas more advance learners who have developed high prior knowledge on the topic benefit more from problem-solving practice (Kalyuga et al. 2003; Kalyuga et al. 2001; Kalyuga 2007; Reisslein et al. 2006).[1]

Some research has shown that the transition from studying worked examples to problem-solving practice can be facilitated through intermediary approaches in which learners are given partially worked out solutions and have to complete some of the solution steps (van Merriënboer et al. 2003), or through a faded example procedure in which the solution steps for fully worked examples are presented and then increasingly removed on subsequent exposures (Renkl and Atkinson 2003; Renkl 2014). The idea behind these approaches is that the incremental introduction to problem solving helps learners practice applying the solution procedures without overtaxing the cognitive resources they need to manage the demands associated with independent problem solving.

Thus, the degree to which novice learners need the extra scaffolding afforded by worked examples may depend on the complexity of the task. Consistent with this notion, it has been proposed that the benefits of worked examples increase as element interactivity increases (Leahy et al. 2015; van Gog and Sweller 2015; van Gog et al. 2015). This idea was directly explored recently by Hanham et al. (2017), who compared the effects of retrieval vs. studying of worked examples on students' writing skills under different levels of element interactivity. Students first received a worked example to illustrate the objective of the task (e.g., writing a persuasive argument or a poem), and then were provided with an additional worked example or attempted to write the material on their own. A final test over the relevant writing skills revealed that students benefited by trying to write the material on their own instead of studying extra worked examples under conditions in which element interactivity was low (i.e., a template was provided for the persuasive argument in Experiments 1 and 2), but this advantage was eliminated under conditions in which element interactivity was increased (i.e., by requiring students to learn several rules for constructing a particular type of poem in Experiment 3), and significantly reversed under conditions that further increased element interactivity (i.e., by using the same poem task but including younger students as participants in Experiment 4, and removing the instructional guidance that illustrated the rules in Experiment 5 (although a replication with a delayed final test yielded no performance differences between groups)).

This research suggests that the benefits of retrieval are not ubiquitous and may depend on the complexity of the task. Much of the research on retrieval practice with novice learners has been conducted with declarative memory tasks for which the level of element interactivity is low. Benefits of retrieval are regularly observed under these conditions, but have not been observed under conditions involving principle-based rules or problem-solving procedures when the level of element interactivity is high. From a cognitive load perspective, studying

[1] The order of participants' choices indicated that they sometimes used the restudy opportunities as feedback following the test opportunities, particularly for definitions they could not recall.

worked examples in problem-solving tasks reduces the extraneous load that would otherwise occur by trying to solve the problems using inefficient strategies that are not based on knowledge of the correct procedures (e.g., trial-and-error or means-end analysis), and this reduced extraneous load frees up resources to focus on the relevant problem-solving procedures that are illustrated by the examples and that contribute to the development of a solution-based schema.

In a recent re-conceptualization of cognitive load theory, Kalyuga and Singh (2016) proposed that the development of a solution-based schema may not be the only goal of learning in a complex task, however. In problem-solving tasks, learners need to induce underlying rules and procedures associated with the solution, but they also need to successfully remember those rules and procedures for future use. Effective learning depends on the optimal managing of cognitive resources in a way that facilitates the goals of the task, and those goals might entail different processing and thus different approaches to effectively managing those resources. In this way, depending on the goals for learning (i.e., rule induction vs. memory proficiency), a given manipulation may produce different outcomes even within the context of the same task. There is much evidence supporting the use of repeated studying (i.e., worked examples) in facilitating rule induction (Atkinson et al. 2000; Baars et al. 2014; van Gog et al. 2011; van Gog and Kester 2012; van Gog et al. 2015); however, if retrieval more effectively enhances memory retention than study, then memory for the rule itself might benefit from retrieval. To the extent that enhanced memory for the rule is needed for future problem-solving performance (i.e., after a delay), retrieval could be more effective than study for facilitating future problem solving.

In the only known study to investigate this within the same task context, Yeo and Fazio (2019) compared studying of worked examples vs. problem-solving practice using non-identical isomorphic problem scenarios (the typical procedure in studies of worked examples) vs. identical problem scenarios in which the probability problem was the same from one trial to the next but the specific numbers were altered. On 1-week delayed final tests containing new problems, non-identical problem scenarios produced a benefit of worked examples over problem-solving (Experiment 1), but identical problem scenarios produced a benefit of problem-solving over worked examples (Experiment 2). Thus, even in the context of the same problem-solving task, the effectiveness of retrieval over study depended on the goals for learning. Study was more effective than retrieval for inducing the solution procedure from non-identical problems, whereas retrieval was more effective than study for remembering the solution procedure that was practiced with identical problems. The authors concluded that both study and retrieval can benefit learning in problem-solving tasks, but they do so through different mechanisms. These data are consistent with those from studies on deductive inferences and analogical reasoning tasks showing that the benefits of retrieval depend upon the degree to which the ultimate learning outcomes depend on enhanced memory for the information needed to solve the problem (Eglington and Kang 2018; Hostetter et al. 2019; Peterson and Wissman 2018; Wissman et al. 2018). Given the contribution of memory to higher-order problem solving and transfer performance (Barnett and Ceci 2002), it appears that retrieval can benefit learning in such higher-order tasks through enhancing memory for the principles and procedures that must be applied.

Retrieval Practice in Self-Regulated Learning

Although empirical research can provide much evidence about the effectiveness of particular learning techniques, those techniques are not likely to benefit learning if students do not use them.

Therefore, it is important to explore whether students utilize retrieval in their own studying, and whether they utilize it in the ways that would be most effective. The results of the studies reviewed above suggest that retrieval is beneficial for improving learning in tasks that rely primarily on memory recall, but not always in more complex tasks that rely on the induction of problem-solving rules and procedures. When learning information, do students make decisions consistent with this evidence and choose to engage in retrieval when it would be most effective for their learning?

This question has been addressed in a number of studies that allow participants to select their own strategies while studying information. We divide this section into studies that have explored students' self-regulated learning in memory-based tasks and in more complex problem-solving tasks. Both sections include measures of students' self-regulated learning decisions involving retrieval in both laboratory- and classroom-based environments.

Memory-Based Tasks Studies exploring memory-based learning (which should benefit from retrieval) show that students commonly avoid using retrieval in their studying. In a study by Karpicke (2009, Experiment 4), participants learned a list of Swahili-English word pairs for an upcoming memory test 1 week later. After an initial encoding phase over each word pair, participants were asked to make a decision about whether they would like to (a) test their memory of that word pair, (b) study the word pair again, or (c) drop the word pair from the list and not see it again. Despite evidence that attempting retrieval would be the most effective strategy (based on the results of a preliminary study using the same items), participants chose more often to restudy the word pairs (about 55% of items) than to be tested over them (about 35% of items).

A similar preference for restudying over testing was found by Dunlosky and Rawson (2015a, Experiments 1-2) in a study where participants learned definitions for key terms from developmental psychology (e.g., object permanence). After studying each term and definition once, participants were given unlimited opportunities to learn the terms and definitions by (a) restudying each term and definition together, (b) trying to recall the definition for each term, or (c) judging the accuracy of their recall using feedback of the correct definition for each term. Participants could choose any of the three options for any of the terms, in any order, as often as they liked. Results showed that participants engaged in more overall practice trials involving restudy (choosing to restudy each item about 2.6 times, on average) than testing (choosing to recall each definition less than 2 times, on average), and they chose restudying and testing more often than judging the accuracy of their answers (which they chose about 1.4 times per item, on average). When given the option, participants thus appear to choose restudy opportunities more often than retrieval opportunities in their self-regulated learning.[2]

Further research has revealed that students' choices of when to use retrieval for memory-based tasks depend on other factors. In particular, the difficulty of the learning task can influence these decisions. Tullis, Fiechter, and Benjamin (2018, Experiments 2-5) presented participants with word pairs that were either related (easier to learn) or unrelated (harder to learn), and following each word pair participants were allowed to choose whether they wanted to (a) see that word pair again for restudy, (b) be tested over that word pair, or (c) drop that

[2] Similar interactions have been reported in memory-based tasks, where learners with high prior knowledge show larger benefits of retrieval over restudy (Carpenter et al., 2016), and learners with higher fluid intelligence show larger benefits of retrieval for difficult items compared to easier items whereas learners with lower fluid intelligence show larger benefits of retrieval for easier items than for difficult items (Minear, Coane, Boland, Cooney, & Albat, 2018).

word pair so that it would not be seen again. Participants' choices of which strategy to use depended upon the type of item, such that for unrelated items participants more often chose to restudy than to be tested, and for related items they more often chose to be tested than to restudy. Thus, participants chose testing more often for easier items and restudy more often for more difficult items. During the practice phase, half of participants' choices were honored (i.e., they were given a restudy opportunity when they chose restudy, and a test opportunity when they chose test), and half were dishonored (i.e., they were given a restudy opportunity when they chose test, and a test opportunity when they chose restudy). When practice tests were accompanied by feedback (Experiments 4–5), a final test 2 days later revealed an overall benefit of testing over restudying regardless of whether participants' choices had been honored or dishonored. That is, despite participants' decisions to use testing for easier items more than for difficult items, testing with feedback ended up resulting in general learning benefits over restudying for both types of items.

Similar results were observed by Toppino et al. (2018), who found that participants were more likely to choose to restudy word pairs under more difficult learning conditions (i.e., when word pairs were unrelated and a longer lag was inserted in-between initial study and subsequent presentation of each item) and more likely to choose testing under easier learning conditions (i.e., when word pairs were related and a shorter lag was inserted in-between initial study and subsequent presentation). The results of these two studies indicate that students seem to rely on the perceived difficulty of a retrieval opportunity when making their decisions about whether or not to use retrieval in their self-regulated learning. Students' reliance on perceived difficulty or effort as a cue for self-regulated learning could be misleading, however, as Tullis et al. (2018) found that testing with feedback—regardless of item difficulty—produced better retention than restudy. Thus, students may misinterpret the role of effort in self-regulated learning, such that they avoid choosing strategies that involve greater effort, even when those strategies are objectively better for learning.

Indeed, the perceived effort involved in retrieval appears to be a key factor underlying students' decisions to use it in their self-regulated learning. Kirk-Johnson, Galla, and Fraundorf (2019, Experiment 2) had participants learn two text passages, one through retrieval (i.e., answering fill-in-the-blank questions from the passage after reading it) and the other through restudy (i.e., reviewing intact statements from the passage after reading it). Immediately after each, participants were asked to rate the mental effort involved in the task and how much they thought they had learned from it. After reading both passages and experiencing both retrieval and restudy, participants answered further questions about which strategy required more effort and which one was more effective for their learning. Participants were then given a third passage to read and were asked to choose whether they would like to review the passage (for an upcoming test) using either the retrieval or restudy strategy. Most participants (69%) chose restudy and a much smaller percentage (31%) chose retrieval. Participants also rated retrieval as more effortful, and less effective for their learning, compared to restudy. Of primary interest, these judgments influenced strategy decisions in a particular way. Greater judgments of effort associated with retrieval (compared to restudy) led to a reduced tendency to choose retrieval during self-regulated learning of the third reading passage. This relationship was mediated by the perceived learning effectiveness associated with retrieval, in that greater effort involved in retrieval predicted lower perceived effectiveness of retrieval for learning, which in turn predicted the decision to avoid retrieval and opt for restudy during self-regulated learning. Thus, students' decisions to forego a more cognitively demanding retrieval strategy (compared to an easier restudy strategy) appear to be driven by the assumption that retrieval is not an effective strategy for learning.

These results are consistent with other literature showing a negative relationship between effort and perceived learning (see Baars et al. 2020), and with the results of a number of studies showing that students are largely unaware of the direct benefits of retrieval. In recent survey studies inquiring about undergraduate students' study methods, a fairly large portion of students—over 70%—reported that they regularly test themselves with questions or practice problems (Geller et al. 2018; Hartwig and Dunlosky 2012). In these studies, over 60% of students also reported that they regularly engage in re-reading of course material as part of their studying. When students were asked *why* they use self-testing, the majority of students from these and other survey studies (e.g., see Kornell and Bjork 2007) reported that they use self-testing to figure out how well they have learned the information they are studying. A small portion of students (less than 30%) reported that they use self-testing because they learn better that way than through re-reading, and an even smaller portion (less than 15%) reported that they use self-testing because they enjoy it more than re-reading. Thus, students appear to use self-testing, but primarily as a means of checking their knowledge, not because they feel it helps them learn, and certainly not because they enjoy it.

Such factors may underlie the tendencies of students to underutilize retrieval practice as a means of learning information in their courses. Observational studies tracking students' use of practice questions in their courses confirm that students do not make full use of retrieval-based tools for learning. In these studies, students were offered one or more opportunities to complete practice quizzes prior to their mandatory exams. The practice quizzes were completely optional, and no incentives were provided for completing them. Carillo-de-la-Peña et al. (2009) investigated students' completion rates of an optional practice quiz across different courses that included basic and advanced psychology, genetics, ecology, zoology, and medicine. The completion rates varied among the courses, such that the optional quiz was completed by about 40% of basic psychology students, 50% of advanced psychology students, 60% of genetics students, and 66% of ecology students. The completion rates of the optional quiz for students in the zoology and medicine courses were much higher, at 92% and 91%, respectively. Other studies using a similar type of optional quiz prior to a mandatory exam have observed completion rates of about 46% for dentistry students (Olson and McDonald 2004), and about 75% for pre-med students (Velan et al. 2008).

In the case of multiple opportunities to complete optional practice quizzes, the available data indicate that students might use these tools even less often. Throughout the course of an entire academic term, Johnson (2006) provided two optional online practice quizzes over each textbook chapter in an educational psychology course. Each of the 28 quizzes (two quizzes per chapter covering 14 chapters) was available to students from the first day of class until the end of final exams. Although no credit or other external incentives were provided for completing these quizzes, students were informed of the potential benefits of the quizzes as a study aid and periodically reminded about them throughout the term. Although nearly 70% of students completed at least one quiz throughout the term, the overall completion rate of the quizzes was quite low. Of the 28 quizzes available, on average students completed only about three of them, or approximately 20% of the available quizzes. Furthermore, fewer than 10% of students completed all of the 28 quizzes.

Perhaps students want to review course material but prefer not to use a quiz format because, as indicated by the results of several survey studies, they do not consider quizzes effective learning tools, or they do not find quizzes enjoyable (e.g., Geller et al. 2018; Hagemeier and Mason 2011; Hartwig and Dunlosky 2012; Kornell and Bjork 2007; McAndrew et al. 2016). If students were given opportunities to review material by simply reading it (rather than answering quiz questions), would they more likely engage in online practice questions?

A recent study by Carpenter et al. (2017) explored this question. Students in an introductory biology course were given optional online reviews throughout the academic term. Several review questions pertaining to each day's lesson were posted online after each class and were available for students to access for the remainder of the term. Students could access these reviews in either *test* format (consisting of multiple-choice questions followed by feedback of the correct answers) or in *read* format (consisting of the same multiple-choice questions, but with the correct answers highlighted so that students did not need to answer the questions themselves). Both test and read formats were available, and students could access the questions using either format, or both formats, as often as they liked. Students were informed by the instructor that the reviews were available, and students were periodically reminded about the reviews, but completion of the reviews was entirely optional and was not incentivized. Under these conditions, just over half of the students completed at least one review prior to the exams over that material. Most students used the test format or some combination of test and read formats (about 50% of students) compared to the read format alone (only about 5% of students). Similar to Johnson's (2006) study, completion rate of the entire set of reviews was fairly low, with only about 30% of students completing all of the available reviews prior to the exams over that material.

Using a similar type of approach in an introductory psychology course, Corral et al. (2020) found that students more often utilized the test format over the read format in completing optional online reviews, but fewer than 15% of the total number of online reviews were completed. As part of the study, a survey was included at the end of the term inquiring about students' reasons for using—and *not* using—the online reviews. Of the students who used the online reviews, the most common reason for doing so (chosen by 55% of students) was that students felt that the review questions helped them figure out how well they knew the information. Of the students who did *not* use the reviews, the most common reason for avoiding the reviews (chosen by 40% of students) was that students used other methods of studying and felt that they did not need to use the reviews. Further analysis of the timing of completion of the reviews revealed that, although the reviews were available immediately after each class, students tended to delay the completion of the reviews until shortly before the exams, which further indicates that students may have been using the reviews as a check of their knowledge in the final stages of studying. These findings are consistent with the results of the earlier survey studies (Geller et al. 2018; Hartwig and Dunlosky 2012), in that students tend to view practice questions primarily as a tool for checking their knowledge rather than for enhancing their learning, and they further suggest that a lack of awareness about the direct benefits of practice questions may lead students to avoid using them as a study strategy.

Problem-Solving Tasks What about more complex tasks? For novice learners, retrieval (i.e., problem-solving practice) has been found to be ineffective, and sometimes counter-effective, compared to studying worked examples (Atkinson et al. 2000; Baars et al. 2014; van Gog et al. 2011; van Gog and Kester 2012; van Gog et al. 2015; Yeo & Fazio, Experiment 1). Benefits of retrieval have been observed in learning situations that involve identical examples (Yeo and Fazio 2019, Experiment 2), however all of the known studies on self-regulated learning using problem-solving vs. worked examples have involved non-identical examples during learning. Given the tendency for worked examples to be as effective, or more effective, than retrieval under these conditions, optimal self-regulated decisions would presumably involve frequent use of worked examples during study.

However, based on a handful of studies it appears that novice learners do not typically prefer worked examples when given the option. Foster, Rawson, and Dunlosky (2018, Experiment 1) presented novice learners with probability problem scenarios in the form of problems to solve vs. worked examples alternating with problems, or they allowed students to choose whether they wanted to receive a worked example or a problem on each of several trials (i.e., the self-regulated group). On a final test over new problems representing the same principles, students who received alternating worked examples/problems significantly outperformed those who received only problems, replicating the known benefits of worked examples. Participants in the self-regulated group chose worked examples only 27% of the time and performed only slightly better on the final test than students who received problems with no worked examples. When comparing faded examples to worked examples, faded examples were as effective (Experiment 2) or more effective (Experiment 3) than worked examples, and a self-regulated group that had the option to select partially worked examples (with the final solution step omitted) did so almost as often as they selected problem solving (choosing problem solving and partially worked examples about 40% of the time, and choosing worked examples only about 15% of the time), ultimately performing at a similar level to the faded example group on the final test.

Why did students underutilize worked examples? Further analyses by Foster et al. (2018) revealed that worked examples and partially worked examples were most often chosen after a failed problem solving attempt. Although students received feedback after incorrect problem solving attempts, the feedback provided a general description of how to solve the problem and did not provide the same detailed step-by-step procedure that was present in the worked examples. It appears, therefore, that students may have chosen problem solving practice as a means of checking their understanding of the problem solutions, and then utilized the worked examples as opportunities to view detailed information about the procedures in cases where they could not solve the problem successfully. Furthermore, a questionnaire after Experiment 3 revealed that students perceived problem solving as more difficult, and less effective for learning, than worked examples or partially worked examples. This finding is consistent with self-reported data from other studies of worked examples, in which problem solving is often perceived as less effective for learning than worked examples (Baars et al. 2014; Yeo and Fazio 2019).

Students' choices to use problem solving more often (despite considering it less effective than worked examples) appear to reflect a mismatch between which method helps them learn and which method they choose. Why do students *not* more often choose the method they believe is better for learning? The answer may lie in what students perceive as the primary purpose of retrieval. If retrieval is viewed primarily as a means of checking their understanding and subsequent worked examples function to clarify and remediate aspects of the procedure that are not well understood, students may be engaging in retrieval more often (to check understanding), but still believe that the learning primarily takes place during the studying of worked examples. Particularly if students are trying to master the procedures, frequent checking of their understanding provides them with information about their progress, allowing them to know when they have reached this goal.

Indeed, the desire for mastery could underlie students' study decisions regarding worked examples in real courses. Tempelaar et al. (2020) explored students' engagement with practice problems in an introductory mathematics course. To prepare for mandatory quizzes, students used a computerized program that provided practice problems illustrating different mathematical principles. For each problem, students had the option of attempting to solve it, asking for a hint, or asking for the fully worked example. They could engage with any of the problems as often as they wished, and if they solved an adequate number of the problems correctly they could earn bonus credit on the quizzes. Students were thus incentivized to solve as many problems correctly as they could. Of the total number of problems that students engaged with, on average almost 70% of the time they attempted to solve the problem and about 30% of the time they chose to view the worked example (hints were seldom ever chosen). Although timing data were not reported to allow an analysis of when students relied on worked examples at the individual problem level (e.g., after failed problem solving attempts, as in Foster et al. 2018), it appears that students exhibited a preference for problem-solving under conditions where they were incentivized to demonstrate mastery of the material by solving the problems. It is worth noting, however, that the way in which the practice problems were set up (initially in problem format, with the option to select the worked example after seeing the problem) could have encouraged students to attempt problems more often and reserve the worked example option for those problems that they could not solve or felt unsure about. As such, potential differences in problem difficulty and performance incentives favoring problem-solving practice could have influenced students' tendencies to opt for problem-solving practice.

Lui et al. (2019) found that students relied heavily on problem-solving practice to learn computer algorithms in an operating systems course. As a supplement to regular course activities, students were given access to a computer program that allowed them to study algorithms by choosing problem solving exercises and then checking their answers, or by skipping the problem and going right to the solution in the form of a worked example. The program was optional and students could engage with it at any time, using either of the formats. Of the problems that students engaged with, they attempted to solve the problem over 96% of the time, and only about 4% of the time skipped the problem to view the worked example. These findings suggest a heavy preference for problem-solving practice over worked examples. However, it is important to note that the optional program was utilized by only about half of the class, which could have represented a select group of students who preferred to use problem-solving practice. The program also provided detailed feedback of the solution steps after attempting a problem (unlike Foster et al. 2018, and Tempelaar et al. 2020, in which these detailed steps were not provided as feedback). If students use worked examples primarily to check their understanding of the solutions to problems, the detailed feedback offered after the problem attempt could have reduced students' tendencies to rely on the worked examples for feedback. Furthermore, analysis of students' engagement with the problem-solving exercises over time showed a dramatic increase in the use of the exercises right before scheduled quizzes in the course, followed by an abrupt drop in the use of the exercises right after the quizzes, which might suggest (similar to the findings of Corral et al. 2020 on the use of optional practice questions in courses) that students utilized the problem-solving exercises to check their understanding and level of preparation for the quizzes.

Although research on self-regulated learning of retrieval in problem-solving tasks is limited, the findings so far suggest that students have a preference for choosing problem-solving practice over studying worked examples. Data from real courses are complex and could be influenced by performance incentives and a number of characteristics unique to individual students and the particular learning task (e.g., see van Gog et al. 2020). These data also do not demonstrate whether or not students' decisions were optimal, as individual students' tendencies to engage with problem-solving or worked example practice were not linked with later learning outcomes. Though further research from real courses is clearly needed, findings on the timing of the use of worked examples at the individual student level (from Foster et al. 2018) suggest the possibility that students use problem-solving practice primarily as a means of checking their understanding of problem solutions, and relying on worked examples to remediate understanding of the procedures after failed problem solving attempts. This parallels the findings on students' self-reported use of retrieval practice, in which students rely on retrieval more as a means of checking their knowledge than as a means of directly learning the material (Geller et al. 2018; Hartwig and Dunlosky 2012).

Effort Involved in Retrieval as a Cue in Self-Regulated Learning

Students perceive that retrieval involves more effort than study (Foster et al. 2018; Hanham et al. 2017; Kirk-Johnson et al. 2019; van Gog et al. 2011). However, students do not appear to be aware of when this effort is good for learning and when it is not good for learning. Retrieval is typically beneficial for memory-based tasks, but not for problem-solving tasks involving non-identical problems during early stages of learning. Students' perceptions of effective learning are generally misaligned with this evidence, however, as they routinely rate retrieval as *less* effective for learning than rereading (Kirk-Johnson et al. 2019; Tullis et al. 2013; Yeo and Fazio 2019) or studying worked examples (Baars et al. 2014; Foster et al. 2018; Yeo and Fazio 2019).

Students' self-regulated learning decisions also appear to be misaligned with the effects of retrieval. Students often choose to restudy information instead of retrieve information in memory-based tasks where retrieval would be more beneficial (Karpicke 2009; Tullis et al. 2018), and they often choose to solve problems instead of study worked examples in problem-solving tasks involving non-identical problems where worked examples would be equally or more beneficial (Foster et al. 2018; Lui et al. 2019).

What does this suggest about the factors driving students' study decisions? Students do not consistently choose to engage in the more effortful task of retrieval, but nor do they consistently avoid it. Thus, a pure *effort-avoidance* strategy seems unlikely. Students do not consistently choose the tasks that they consider to be more effective for learning either, which discounts a pure *learning effectiveness* strategy underlying their self-regulated decisions. Although students perceive that retrieval is more effortful, it appears that they may not consider this effort as a direct contributor to learning. The common finding that students rate rereading and worked examples as more effective for learning suggests that students might view these as the primary means by which learning takes place. Consistent with self-report survey data, they may view retrieval more as a way of *revealing* that learning than as a direct strategy for *affecting* learning (Geller et al. 2018; Hartwig and Dunlosky 2012).

When viewed in this way, students' self-regulated learning decisions may reflect the use of retrieval primarily as a monitoring strategy. Retrieval effectively distinguishes learned material (that which could be retrieved) from unlearned material (that which could not be retrieved), allowing learners to focus subsequent study efforts on the unlearned material. From a *discrepancy-reduction* framework of self-regulated learning (Dunlosky and Ariel 2011), retrieval may serve as the means by which the discrepancy between students' current state of knowledge and their goal state is revealed, which triggers self-regulated decisions aimed at reducing that discrepancy. To the extent that study opportunities are viewed as the primary means by which learning takes place, students' efforts to reduce this discrepancy may involve the choice to engage in further study in order to increase learning, and to reserve retrieval for those times when they want to check what they have learned.

The effort involved in retrieval may provide further diagnostic information that is used to monitor learning. That is, in the process of checking what they have learned, students might rely on the effort they experience during retrieval as a reflection of how well they have learned something from a previous study opportunity. A retrieval experience that is successful and relatively fluent might be interpreted as a sign that the previous study opportunity was effective, whereas a more difficult and less fluent retrieval experience (even if successful) might be interpreted as a sign that the study opportunity was less effective. Indeed, it has been shown that students assign higher judgments of learning to items that they could recall more quickly (e.g., Koriat and Ma'ayan 2005), suggesting that the ease or fluency with which an answer is retrieved serves as an important cue for students' impressions about how well they have learned something. A difficult retrieval experience (reflecting less effective learning) may thus lead students to engage in further study during subsequent learning opportunities as a means of improving learning for that material.

Thus, the choice to engage in study (instead of retrieval) could be driven by different things. When students view retrieval as a learning strategy, they could interpret the effort experienced during retrieval as a sign that learning via retrieval was ineffective, and subsequently choose to avoid using retrieval again as a learning strategy (Kirk-Johnson et al. 2019). When students view retrieval as a monitoring strategy, however, they could interpret the effort experienced during retrieval as a sign that the initial study opportunity was ineffective, and subsequently choose to engage in study because they believe that is the means by which learning takes place. This raises an interesting question about whether retrieval effort drives students to *avoid* things (i.e., no longer use retrieval) or to *pursue* things (i.e., seek more study opportunities). Although the outcome is the same—i.e., choose study over retrieval—the reasons for this may be different depending on students' views of the purpose of retrieval.

Research on students' use of retrieval in self-regulated learning is still in its early stages. Additional data are needed on when, how, and why students engage with retrieval during learning, and how these decisions affect learning outcomes in different types of tasks. Given the inherent effort involved in retrieval, it is particularly important to measure this effort in valid and reliable ways (see Scheiter et al. 2020), understand how the diagnosticity of effort as a cue may be influenced by additional factors such as students' prior knowledge and task design (see van de Pol et al. 2020), and understand factors that influence the sustained investment of effort during prolonged study episodes (see Eitel et al. 2020). Key objectives for future research include the following: (1) collecting observational measures of students' decisions to use retrieval in real learning situations, (2) direct querying of students' reasons for the use of retrieval, (3) incorporating multiple measures of effort during retrieval (e.g., subjective ratings of effort or fluency, or time to complete learning activities) and analyzing how these measures coincide with perceived effectiveness of learning from retrieval, (4) tracking students' responses to effortful retrieval situations, and the actions they take (e.g., persistence, disengagement) to manage that effort, and (5) how such measures affect learning outcomes and are influenced by characteristics of individual learners, type of learning material, and goals for learning.

Encouraging the Optimal Use of Retrieval in Self-Regulated Learning

Understanding the factors underlying students' use of retrieval, and how they influence learning outcomes, can help in formulating ways to optimally use retrieval for learning. To the extent that students do not view retrieval as a learning opportunity, directly informing them about the effects of retrieval in different learning situations could lead students to adopt it as a deliberate strategy for learning. Indeed, learners can be made aware of the benefits of retrieval practice for learning declarative factual information (Tullis et al. 2013), and there is evidence that increased awareness of these benefits results in decisions to choose retrieval more often in self-regulated learning (Ariel and Karpicke 2018; Dunlosky and Rawson 2015a).

When the goal is to enhance memory proficiency, an increased awareness of the benefits of retrieval practice might be particularly effective for encouraging its use, particularly in real courses where students would presumably be motivated to learn the material. There is some evidence that this approach can be efficacious. For example, Kibble (2011) provided optional online quizzes prior to mandatory exams in an undergraduate physiology course. To encourage students to use the quizzes, students were given a presentation on the benefits of quizzes for formative assessment purposes. Students were also shown correlational data from a previous class on the positive relationship between use of the quizzes and subsequent exam scores. Under these conditions, 88% of the students used the quizzes, compared to 52% of students from a previous class that was not given the presentation over the potential benefits of the quizzes (Kibble 2007).

Carpenter et al. (2017) further found that biology students' use of optional online practice questions significantly increased following an instructor-led presentation of the potential benefits of the practice questions. Students' voluntary use of the practice questions during the first half of the course indicated limited use of the questions, with about half of the students trying out at least one set of practice questions before the exam, and only about 30% of students completing all of the practice questions available before the exam. Prior to the exams during the second half of the course, the instructor showed students a brief in-class presentation to remind them about the practice questions, and presented the averaged anonymized exam scores associated with students who had used the practice questions during the first half of the course, which were significantly higher than for students who had not used the practice questions. Following this presentation, students' use of the practice questions significantly increased such that nearly 80% of the class completed at least some practice questions prior to the next exam, and nearly half of the class completed all of the practice questions. Students who completed the practice questions scored significantly higher on exams compared to students who did not complete them, and the number of times students completed the practice questions correlated positively and significantly with their subsequent exam scores. Although data from these studies cannot be interpreted as cause-and-effect relationships because students were free to choose whether and how often to use the practice questions, the pattern of results is in line with those of more controlled experimental studies showing that students' learning benefits from repeated opportunities to engage in retrieval practice (Carpenter et al. 2008; Dunlosky and Rawson 2015a; Kuo and Hirshman 1996; Rawson and Dunlosky 2011).

Although these results are encouraging, interventions that aim to increase awareness of the effects of retrieval may not always work. Even if students are informed about the effectiveness of particular learning techniques, students may not always alter their perceptions or beliefs about those techniques (Yan et al. 2016). Further, although students may adopt effective learning techniques in the short-term after being informed about them, there is evidence that they revert back to their original strategies and habits over time (Susser and McCabe 2013). In this way, effective learning techniques might be viewed in the same way that people sometimes view health-conscious behaviors. One can be fully aware that a nutritious diet and exercise are good, but choose not to do these things because the perceived cost required to implement them (relative to more convenient options) is too great in the immediate short-term. Therefore, interventions aimed at improving students' awareness of effective learning strategies may not be enough to alone produce sizeable or lasting changes in their self-regulated learning.

Alternative intervention approaches might therefore involve other types of incentives that can promote the use of retrieval for enhancing memory. For example, Vaughn and Kornell (2019, Experiment 1) found that participants who were studying word pairs (e.g., idea: seeker) were more likely to choose retrieval over restudying when they could receive hints about the target word they were trying to recall. Participants could choose, for example, to retrieve the target word without any hints (e.g., idea: _____), to retrieve the target word with some of the letters filled in (e.g., idea: se__er), or to restudy the intact word pair. Consistent with past research (e.g., Karpicke 2009), participants chose to study the intact word pair more often (29% of items, on average) than to retrieve it with no hints (only 2% of items, on average). However, the most popular choice by far was to retrieve with hints (69% of items, on average). A subsequent experiment (Experiment 3) showed that when participants were randomly assigned to learn the word pairs through retrieval without hints, retrieval with hints, or restudy, a final test over all of the word pairs revealed a significant advantage of retrieval, regardless of whether or not hints were provided. In other words, retrieval of *any* kind benefited learning relative to restudy. These results are encouraging in that providing hints during retrieval appears to significantly increase participants' desire to choose retrieval, without any apparent costs.

Currently it is unknown why hints increase the desire to use retrieval while studying. Vaughn and Kornell (2019) proposed that hints could activate people's natural desire to get an answer correct, enticing them to attempt retrieval under conditions where they are more likely to be successful than when they may have no idea of what the answer is. In this way, hints could incentivize retrieval because of a positive affective experience associated with retrieving correctly. Along similar lines, hints could increase participants' curiosity to know the answer. Previous research has shown that participants learn better under conditions in which they are highly curious to know the answer to something (e.g., Berlyne 1954; Kang et al. 2009), and that conditions that increase or prolong this curiosity (such as brief delays before the answer is presented) can enhance learning even more (Mullaney et al. 2014). Hints might activate learners' curiosity by providing them with some clues about the answer—but not completely short-circuiting the retrieval process—that still benefits memory but also enhances the motivational or affective experiences associated with retrieval.

Students might know that retrieval is good but have a hard time implementing it because effective ways to use retrieval are neither intuitive nor easy. It might make sense that retrieving something multiple times is better than once, but how often and when should repeated retrieval be done? In the vast literature showing benefits of retrieval on memory retention, few studies have yielded clear prescriptive recommendations for effective ways of using repeated retrieval that are also straightforward for students to implement in their own studying. One such approach—the method of *successive relearning*—involves cycling through a list of to-be-learned material (e.g., terms and definitions) using retrieval with feedback, until all of the material can be recalled with 100% accuracy. Some days later, the same material is retrieved again, followed by feedback, until all of it can be recalled to 100% accuracy again. The same process repeats over several sessions, each time with some of the information being forgotten between sessions but re-learned to 100% accuracy. Rawson et al. (2013) found that this approach, compared to the same number of sessions involving pure study, significantly enhanced long-term memory retention of terms and definitions from an undergraduate psychology course. Rawson and Dunlosky (2011) further found that learning key term definitions to a criterion of three correct recalls during an initial learning session (compared to only one correct recall) required only a fraction of the attempts to re-learn the same material to 100% accuracy in subsequent sessions, and produced significantly better long-term retention. Thus, the practical advice offered from this research is that students should engage in recall-with-feedback practice until they can correctly recall information three times, and then successively re-learn the information in later sessions at least a few days apart. Though research is still limited on the use of such techniques by students, there is evidence that providing students with clear and prescriptive guidance for successive relearning can be quite effective for increasing their tendencies to use it (Dunlosky and Rawson 2015b). Given the effectiveness of such techniques for learning information that lends itself to direct recall that can be checked with feedback (e.g., terms and definitions, vocabulary), future research should explore additional ways that communicating these techniques to students might increase their decisions to use such techniques in their studying.

In more complex forms of learning, optimal use of retrieval may depend on learner characteristics (e.g., prior knowledge) and the goals for learning. Classroom-based studies have shown that students make use of practice problems when those problems are made available in the form of easy-to-access computerized programs (Lui et al. 2019; Tempelaar et al. 2020). Students' tendencies to choose problem-solving over worked examples may be driven partly by their desire to check their knowledge and identify what they already understand and what they do not understand, possibly in an attempt to increase the efficiency of their study.

There is limited research on improving self-regulated learning in problem-solving tasks, and the optimal approaches may be quite different and more complex than for simple memory-based tasks (e.g., Ackerman and Thompson 2017). Some research has shown positive effects of self-regulation training, in that learning outcomes benefit from teaching students to guide their study decisions based on their performance and degree of mental effort invested in the current task (van Gog et al. 2020). In order to align students' decisions with effective learning for complex problem-solving tasks, it may also be desirable to provide problems with partially worked solutions rather than full problem solving. Experimental studies have demonstrated the benefits of partial problem solving (e.g., Renkl 2014; Renkl and Atkinson 2003; van Merriënboer et al. 2003). Furthermore, in studies where students are given an option to complete partial problems vs. full problems, students select partial problem solving about as often as full problem-solving (Foster et al. 2018, Experiments 2–3), and sometimes substantially *more* often than full problem solving (van Merriënboer et al. 2002), with later benefits on problem-solving performance higher following self-regulated use of partial problem-solving compared to a group that engaged only in full problem-solving. Thus, partial problem solving is effective for learning *and* preferred by students. Though the exact reasons for this preference are not immediately apparent, it is possible that partial problems could offer some of the same appeal as hints (e.g., Vaughn and Kornell 2019) in arousing students' curiosity to know the answer under conditions where they are more likely to be correct. Partial problems might also encourage effective self-regulated learning via performance feedback. That is, compared to full problem solving that can reveal knowledge deficits without always pinpointing where the deficits are, partial problems might more effectively reveal knowledge deficits in a particular solution step that can be targeted with further practice. Because of their tendency to identify which solution procedures are understood, partial problems might also effectively signal to learners (after a high degree of accuracy has been maintained) when they have developed proficiency and are ready for more advanced practice with full problem solving. Because optimal use of problem-solving practice depends upon regular assessments of learners' proficiency, partial problem solving practice could be beneficial not only for building this proficiency but also for determining when learners have reached it.

Conclusions

Retrieval practice has been a widely studied technique that can be highly effective for enhancing memory proficiency, but students' use of retrieval in their self-regulated learning is at present not well-understood. When given the option, students tend to avoid using retrieval in contexts where it would be beneficial (enhancing memory), and they tend to overuse it in contexts where it would be better to start with extended study rather than retrieval (inducing problem-solving rules and procedures). We proposed that students' self-regulated learning decisions involving retrieval may be driven by students' views about the purpose of retrieval. To the extent that students view retrieval as a means of checking their knowledge (rather than as a means of directly affecting learning), they may be inclined to rely on retrieval to gauge how effectively they have learned from previous study opportunities. An effortful retrieval experience may be interpreted as a sign that the previous study opportunity was ineffective, leading students to choose further study opportunities. Alternatively, students may utilize retrieval as a means of enhancing their learning, and the effort involved in retrieval may be interpreted as a sign that the retrieval opportunity itself was ineffective as a learning opportunity, leading them to choose study instead of retrieval when given the option, but this time as a means of directly avoiding retrieval because it is considered an ineffective strategy. Current data on students' study decisions cannot clearly distinguish between these possibilities, raising interesting questions for future research about the intentions underlying students' use of retrieval, the factors influencing those intentions, and the consequences for self-regulated learning decisions and ultimate learning outcomes. Understanding more about how students think about retrieval will contribute new data that informs the EMR framework (de Bruin et al. 2020) by shedding light on how students interpret the source of experienced effort during learning, whether they can identify this effort as productive or non-productive for learning, and whether the steps they take to deal with that effort afford the optimal management of cognitive resources to meet particular learning goals.

Compliance with Ethical Standards

Conflict of Interest Shana Carpenter has received grants from the National Science Foundation (DUE 1504480) and the James S. McDonnell Foundation (220020483).

References

Abott, E. E. (1909). On the analysis of the factor of recall in the learning process. *The Psychological Review: Monograph Supplements, 11*, 159–177.

Ackerman, R., & Thompson, V. A. (2017). Meta-reasoning: Monitoring and control of thinking and reasoning. *Trends in Cognitive Sciences, 21*(8), 607–617.

Adesope, O. O., Trevisan, D. A., & Sundararajan, N. (2017). Rethinking the use of tests: A meta-analysis of practice testing. *Review of Educational Research, 87*(3), 659–701.

Agarwal, P. K., Karpicke, J. D., Kang, S. H. K., Roediger III, H. L., & McDermott, K. B. (2008). Examining the testing effect with open- and closed-book tests. *Applied Cognitive Psychology, 22*(7), 861–876.

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

Arnold, K. M., & McDermott, K. B. (2013). Test-potentiated learning: Distinguishing between direct and indirect effects of tests. *Journal of Experimental Psychology. Learning, Memory, and Cognition, 39*, 940–945.

Atkinson, R. K., Derry, S. J., Renkl, A., & Wortham, D. (2000). Learning from examples: Instructional principles from the worked examples research. *Review of Educational Research, 70*(2), 181–214.
Baars, M., van Gog, T., de Bruin, A., & Paas, F. (2014). Effects of problem solving after worked example study on primary school children's monitoring accuracy. *Applied Cognitive Psychology, 28*(3), 382–391.
Baars, M., Wijnia, L., de Bruin, A., & Paas, F. (2020). The relation between students' effort and monitoring judgments during learning: A meta-analysis. *Educational Psychology Review*.
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.
Berlyne, D. E. (1954). An experimental study of human curiosity. *British Journal of Psychology, 45*, 256–265.
Birnbaum, M. S., Kornell, N., Bjork, E. L., & Bjork, R. A. (2013). Why interleaving enhances inductive learning: The roles of discrimination and retrieval. *Memory & Cognition, 41*(3), 392–402.
Bjork, R. A., Dunlosky, J., & Kornell, N. (2013). Self-regulated learning: Beliefs, techniques, and illusions. *Annual Review of Psychology, 64*(1), 417–444.
Butler, A. C. (2010). Repeated testing produces superior transfer of learning related to repeated studying. *Journal of Experimental Psychology. Learning, Memory, and Cognition, 36*, 1118–1133.
Butler, A. C., & Roediger III, H. L. (2007). Testing improves long-term retention in a simulated classroom setting. *European Journal of Cognitive Psychology, 19*(4-5), 514–527.
Carillo-de-la-Peña, M. T., Baillès, E., Caseras, X., Martinez, À., Ortet, G., & Pérez, J. (2009). Formative assessment and academic achievement in pre-graduate students of health sciences. *Advances in Health Sciences Education, 14*(1), 61–67.
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*, 1563–1569.
Carpenter, S. K. (2011). Semantic information activated during retrieval contributes to later retention: Support for the mediator effectiveness hypothesis of the testing effect. *Journal of Experimental Psychology. Learning, Memory, and Cognition, 37*, 1547–1552.
Carpenter, S. K. (2017). Spacing effects on learning and memory. In J. T. Wixted (Ed.), Cognitive psychology of Memory, Vol. 2 Learning and Memory: A Comprehensive Reference, 2nd edition, J. H. Byrne (Ed.), pp. 465–485. Oxford: Academic Press.
Carpenter, S. K., & DeLosh, E. L. (2005). Application of the testing and spacing effects to name learning. *Applied Cognitive Psychology, 19*(5), 619–636.
Carpenter, S. K., & Yeung, K. L. (2017). The role of mediator strength in learning from retrieval. *Journal of Memory and Language, 92*, 128–141.
Carpenter, S. K., Pashler, H., & Vul, E. (2006). What types of learning are enhanced by a cued recall test? *Psychonomic Bulletin & Review, 13*(5), 826–830.
Carpenter, S. K., Pashler, H., Wixted, J. T., & Vul, E. (2008). The effects of tests on learning and forgetting. *Memory & Cognition, 36*(2), 438–448.
Carpenter, S. K., Pashler, H., & Cepeda, N. J. (2009). Using tests to enhance 8th grade students' retention of U. S. History facts. *Applied Cognitive Psychology, 23*(6), 760–771.
Carpenter, S. K., Lund, T. J. S., Coffman, C. R., Armstrong, P. I., Lamm, M. H., & Reason, R. D. (2016). A classroom study on the relationship between student achievement and retrieval-enhanced learning. *Educational Psychology Review, 28*(2), 353–375.
Carpenter, S. K., Rahman, S., Lund, T. J. S., Armstrong, P. I., Lamm, M. H., Reason, R. D., & Coffman, C. R. (2017). Students' use of optional online reviews and its relationship to summative assessment outcomes in introductory biology. *CBE Life Sciences Education, 16*, ar23.
Cooper, G., & Sweller, J. (1987). Effects of schema acquisition and rule automation on mathematical problem-solving transfer. *Journal of Educational Psychology, 79*(4), 347–362.
Corral, D., Carpenter, S. K., Perkins, K., & Gentile, D. A. (2020). Assessing students' use of optional online lecture reviews. *Applied Cognitive Psychology, 34*(2), 318–329.
de Bruin, A. B. H., & van Gog, T. (2012). Improving self-monitoring and self-regulation: From cognitive psychology to the classroom. *Learning and Instruction, 22*(4), 245–252.
de Bruin, A. B. H., Roelle, J., Baars, M., & EFG-MRE. (2020). Synthesizing cognitive load and self-regulation theory: A theoretical framework and research agenda. Educational Psychology Review.
Delaney, P. F., Verkoeijen, P. P. J. L., & Spirgel, A. (2010). Spacing and testing effects: A deeply critical, lengthy, and at times discursive review of the literature. *Psychology of Learning and Motivation, 53*, 63–147.
Dunlosky, J., & Ariel, R. A. (2011). Self-regulated learning and the allocation of study time. *Psychology of Learning and Motivation, 54*, 103–140.
Dunlosky, J., & Rawson, K. A. (2015a). Do students use testing and feedback while learning? A focus on key concept definitions and learning to criterion. *Learning and Instruction, 39*, 32–44.

Dunlosky, J., & Rawson, K. A. (2015b). Practice tests, spaced practice, and successive relearning: Tips for classroom use and for guiding students' learning. *Scholarship of Teaching and Learning in Psychology, 1*(1), 72–78.
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.
Eglington, L. G., & Kang, S. H. K. (2018). Retrieval practice benefits deductive inference. *Educational Psychology Review, 30*(1), 215–228.
Eitel, A., Endres, T., & Renkl, A. (2020). Self-management as a bridge between cognitive load and self-regulated learning: The illustrative case of seductive details. *Educational Psychology Review*.
Endres, T., Carpenter, S. K., Martin, A., & Renkl, A. (2017). Enhancing learning by retrieval: Enriching free recall with elaborative prompting. *Learning and Instruction, 49*, 13–20.
Foster, N. L., Rawson, K. A., & Dunlosky, J. (2018). Self-regulated learning of principle-based concepts: Do students prefer worked examples, faded examples, or problem solving? *Learning and Instruction, 55*, 124–138.
Fritz, C. O., Morris, P. E., Nolan, D., & Singleton, J. (2007). Expanding retrieval practice: An effective aid to preschool children's learning. *Quarterly Journal of Experimental Psychology, 60*(7), 991–1004.
Gates, A. (1917). Recitation as a factor in memorizing. *Archives of Psychology, 6*, 1–104.
Geller, J., Toftness, A. R., Armstrong, P. I., Carpenter, S. K., Manz, C. L., Coffman, C. R., & Lamm, M. H. (2018). Study strategies and beliefs about learning as a function of academic achievement and achievement goals. *Memory, 26*(5), 683–690.
Gick, M. L., & Holyoak, K. J. (1980). Analogical problem solving. *Cognitive Psychology, 12*(3), 306–355.
Goossens, N. A. M. C., Camp, G., Verkoeijen, P. P. J. L., & Tabbers, H. K. (2014a). The effect of retrieval practice in primary school vocabulary learning. *Applied Cognitive Psychology, 28*(1), 135–142.
Goossens, N. A. M. C., Camp, G., Verkoeijen, P. P. J. L., Tabbers, H. K., & Zwaan, R. A. (2014b). The benefit of retrieval practice over elaborative restudy in primary school vocabulary learning. *Journal of Applied Research in Memory and Cognition, 3*(3), 177–182.
Hagemeier, N. E., & Mason, H. L. (2011). Student pharmacists' perceptions of testing and study strategies. *American Journal of Pharmaceutical Education, 75*(2), 35.
Hanham, J., Leahy, W., & Sweller, J. (2017). Cognitive load theory, element interactivity, and the testing and reverse testing effects. *Applied Cognitive Psychology, 31*(3), 265–280.
Hartwig, M. K., & Dunlosky, J. (2012). Study strategies of college students: Are self-testing and scheduling related to achievement? *Psychonomic Bulletin & Review, 19*(1), 126–134.
Hinze, S. R., Wiley, J., & Pellegrino, J. W. (2013). The importance of constructive comprehension processes in learning from tests. *Journal of Memory and Language, 69*(2), 151–164.
Hostetter, A. B., Penix, E. A., Norman, M. Z., Batsell, W. R., & Carr, T. H. (2019). The role of retrieval practice in memory and analogical problem-solving. *Quarterly Journal of Experimental Psychology, 72*(4), 858–871.
Jaeger, A., Eisenkraemer, R. E., & Stein, L. M. (2015). Test-enhanced learning in third-grade children. *Educational Psychology, 35*(4), 513–521.
Johnson, G. (2006). Optional online quizzes: College student use and relationship to achievement. *Canadian Journal of Learning and Technology, 32*:61.
Kalyuga, S. (2007). Expertise reversal effect and its implications for learner-tailored instruction. *Educational Psychology Review, 19*(4), 509–539.
Kalyuga, S., & Singh, A.-M. (2016). Rethinking the boundaries of cognitive load theory in complex learning. *Educational Psychology Review, 28*(4), 831–852.
Kalyuga, S., Chandler, P., Touvinen, J., & Sweller, J. (2001). When problem solving is superior to studying worked examples. *Journal of Educational Psychology, 93*(3), 579–588.
Kalyuga, S., Ayres, P., Chandler, P., & Sweller, J. (2003). The expertise reversal effect. *Educational Psychologist, 38*(1), 23–31.
Kalyuga, S., Renkl, A., & Paas, F. (2010). Facilitating flexible problem solving: A cognitive load perspective. *Educational Psychology Review, 22*(2), 175–186.
Kang, S. H. K., McDermott, K. B., & Roediger III, H. L. (2007). Test format and corrective feedback modify the effect of testing on long-term retention. *European Journal of Cognitive Psychology, 19*(4-5), 528–558.
Kang, M. J., Hsu, M., Krajbich, I. M., Loewenstein, G., McClure, S. M., Wang, J. T., & Camerer, C. F. (2009). The wick in the candle of learning: Epistemic curiosity activates reward circuitry and enhances memory. *Psychological Science, 20*(8), 963–973.
Karpicke, J. D. (2009). Metacognitive control and strategy selection: Deciding to practice retrieval during learning. *Journal of Experimental Psychology: General, 138*(4), 469–486.

Kibble, J. (2007). Use of unsupervised online quizzes as formative assessment in a medical physiology course: Effects of incentives on student participation and performance. *Advances in Physiology Education, 31*(3), 253–260.

Kibble, J. (2011). Voluntary participation in online formative quizzes is a sensitive predictor of student success. *Advances in Physiology Education, 35*(1), 95–96.

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*, 1–31.

Koriat, A., & Ma'ayan, H. (2005). The effects of encoding fluency and retrieval fluency on judgments of learning. *Journal of Memory and Language, 52*(4), 478–492.

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.

Kuo, T.-M., & Hirshman, E. (1996). Investigations of the testing effect. *American Journal of Psychology, 109*(3), 451–464.

Leahy, W., Hanham, J., & Sweller, J. (2015). High element interactivity information during problem solving may lead to failure to obtain the testing effect. *Educational Psychology Review, 27*(2), 291–304.

Little, J. L., & McDaniel, M. A. (2014). Metamemory monitoring and control following retrieval practice for text. *Memory & Cognition, 43*, 85–98.

Logan, J. M., Castel, A. D., Haber, S., & Viehman, E. J. (2012). Metacognition and the spacing effect: The role of repetition, feedback, and instruction on judgments of learning for massed and spaced rehearsal. *Metacognition and Learning, 7*(3), 175–195.

Lui, A. K.-F., Poon, M. H. M., & Wong, R. M. H. (2019). Automated generators of examples and problems for studying computer algorithms. *Interactive Technology & Smart Education, 16*(3), 204–218.

McAndrew, M., Morrow, C. S., Atiyeh, L., & Pierre, G. C. (2016). Dental student study strategies: Are self-testing and scheduling related to academic performance? *Journal of Dental Education, 80*(5), 542–552.

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., Agarwal, P. K., Huelser, B. J., McDermott, K. B., & Roediger III, H. L. (2011). Test-enhanced learning in a middle school science classroom: The effects of quiz frequency and placement. *Journal of Educational Psychology, 103*(2), 399–414.

McDermott, K. B., Agarwal, P. K., D'Antonio, L., Roediger III, H. L., & McDaniel, M. A. (2014). Both multiple-choice and short-answer quizzes enhance later exam performance in middle and high school classes. *Journal of Experimental Psychology: Applied, 20*(1), 3–21.

Minear, M., Coane, J. H., Boland, S. C., Cooney, L. H., & Albat, M. (2018). The benefits of retrieval practice depend on item difficulty and intelligence. *Journal of Experimental Psychology. Learning, Memory, and Cognition, 44*, 1474–1486.

Mullaney, K. M., Carpenter, S. K., Grotenhuis, C., & Burianek, S. (2014). Waiting for feedback helps if you want to know the answer: The role of curiosity in the delay-of-feedback benefit. *Memory & Cognition, 42*(8), 1273–1284.

Olson, B. L., & McDonald, J. L. (2004). Influence of online formative assessment upon student learning in biomedical science courses. *Journal of Dental Education, 68*(6), 656–659.

Paas, F., Renkl, A., & Sweller, J. (2003). Cognitive load theory and instructional design: Recent developments. *Educational Psychologist, 38*(1), 1–4.

Paas, F., van Gog, T., & Sweller, J. (2010). Cognitive load theory: New conceptualizations, specifications, and integrated research perspectives. *Educational Psychology Review, 22*(2), 115–121.

Pan, S. C., Lovelett, J. T., Phun, V., & Rickard, T. C. (2019). The synergistic benefits of systematic and random interleaving for second language grammar learning. *Journal of Applied Research in Memory and Cognition, 8*(4), 450–462.

Pashler, H., Bain, P. M., Bottge, B. A., Graesser, A., Koedinger, K., McDaniel, M. A., & Metcalfe, J. (2007). Organizing instruction and study to improve student learning. IES Practice Guide. NCER 2007-2004. *National Center for Education Research* http://eric.ed.gov/?id=ED498555. Access date 30 May 2009.

Peterson, D. J., & Wissman, K. T. (2018). The testing effect and analogical problem-solving. *Memory, 26*(10), 1460–1466.

Pyc, M. A., & Rawson, K. A. (2010). Why testing improves memory: Mediator effectiveness hypothesis. *Science, 330*(6002), 335.

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

Rawson, K. A., Dunlosky, J., & Sciartelli, S. M. (2013). The power of successive relearning: Improving performance on course exams and long-term retention. *Educational Psychology Review, 25*(4), 523–548.
Reisslein, J., Atkinson, R. K., Seeling, P., & Reisslein, M. (2006). Encountering the expertise reversal effect with a computer-based environment on electrical circuit analysis. *Learning and Instruction, 16*(2), 92–103.
Renkl, A. (2014). Toward an instructionally oriented theory of example-based learning. *Cognitive Science, 38*(1), 1–37.
Renkl, A., & Atkinson, R. K. (2003). Structuring the transition from example study to problem solving in cognitive skill acquisition: A cognitive load perspective. *Educational Psychologist, 38*(1), 15–22.
Roediger III, H. L., Agarwal, P. K., McDaniel, M. A., & McDermott, K. B. (2011a). Test-enhanced learning in the classroom: Long-term improvements from quizzing. *Journal of Experimental Psychology: Applied, 17*(4), 382–395.
Roediger III, H. L., Putnam, A. L., & Smith, M. A. (2011b). Ten benefits of testing and their applications to educational practice. *Psychology of Learning and Motivation, 55*, 1–36.
Rohrer, D., & Pashler, H. (2010). Recent research on human learning challenges conventional instructional strategies. *Educational Researcher, 39*(5), 406–412.
Rohrer, D., & Taylor, K. (2007). The shuffling of mathematics problems improves learning. *Instructional Science, 35*(6), 481–498.
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.
Salden, R. J. C. M., Koedinger, K. R., Renkl, A., Aleven, V., & McLaren, B. M. (2010). Accounting for beneficial effects of worked examples in tutored problem solving. *Educational Psychology Review, 22*(4), 379–392.
Scheiter, K., Ackerman, R., & Hoogerheide, V. (2020). Looking at mental effort appraisals through a metacognitive lens: Are they biased? *Educational Psychology Review.*
Schwieren, J., Barenberg, J., & Dutke, S. (2017). The testing effect in the psychology classroom: A meta-analytic perspective. *Psychology Learning and Teaching, 16*(2), 179–196.
Susser, J. A., & McCabe, J. (2013). From the lab to the dorm room: Metacognitive awareness and use of spaced study. *Instructional Science, 41*(2), 345–363.
Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. *Cognitive Science, 12*(2), 257–285.
Sweller, J. (1994). Cognitive load theory, learning difficulty, and instructional design. *Learning and Instruction, 4*(4), 295–312.
Sweller, J. (2006). The worked example effect and human cognition. *Learning and Instruction, 16*(2), 165–169.
Sweller, J. (2010). Element interactivity and intrinsic, extraneous, and germane cognitive load. *Educational Psychology Review, 22*(2), 123–138.
Sweller, J., van Merriënboer, J. J. G., & Paas, F. (1998). Cognitive architecture and instructional design. *Educational Psychology Review, 10*(3), 251–296.
Taylor, K., & Rohrer, D. (2010). The effects of interleaved practice. *Applied Cognitive Psychology, 24*(6), 837–848.
Tempelaar, D. T., Rienties, B., & Nguyen, Q. (2020). Individual differences in the preference for worked examples: Lessons from an application of dispositional learning analytics. *Applied Cognitive Psychology, 34(4), 890–905.*
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.
Toppino, T. C., & Gerbier, E. (2014). About practice: Repetition, spacing, and abstraction. *Psychology of Learning and Motivation, 60*, 113–189.
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*(7), 1164–1177.
Tran, R., Rohrer, D., & Pashler, H. (2015). Retrieval practice: The lack of transfer to deductive inferences. *Psychonomic Bulletin & Review, 22*(1), 135–140.
Tullis, J. G., Finley, J. R., & Benjamin, A. S. (2013). Metacognition of the testing effect: Guiding learners to predict the benefits of retrieval. *Memory & Cognition, 41*(3), 429–442.
Tullis, J. G., Fiechter, J. L., & Benjamin, A. S. (2018). The efficacy of learners' testing choices. *Journal of Experimental Psychology. Learning, Memory, and Cognition, 44*, 540–552.
van de Pol, J., van Loon, M., van Gog, T., Braumann, S., & de Bruin, A. (2020). Mapping and drawing to improve students' and teachers' monitoring and regulation of students' learning from text: Current findings and future directions. *Educational Psychology Review.*

van Gog, T., & Kester, L. (2012). A test of the testing effect: Acquiring problem-solving skills from worked examples. *Cognitive Science, 36*(8), 1532–1541.
van Gog, T., & Sweller, J. (2015). Not new, but nearly forgotten: The testing effect decreases or even disappears as the complexity of learning materials increases. *Educational Psychology Review, 27*, 247–264.
van Gog, T., Kester, L., & Paas, F. (2011). Effects of worked examples, example-problem, and problem-example pairs on novices' learning. *Contemporary Educational Psychology, 36*, 212–218.
van Gog, T., Kester, L., Dirkx, K., Hoogerheide, V., Boerboom, J., & Verkoeijen, P. P. J. L. (2015). Testing after worked example study does not enhance delayed problem-solving performance compared to restudy. *Educational Psychology Review, 27*, 265–289.
van Gog, T., Hoogerheide, V., & van Harsel, M. (2020). The role of mental effort in fostering self-regulated learning with problem-solving task. *Educational Psychology Review*.
van Merriënboer, J. J. G., Schuurman, J. G., de Croock, M. B. M., & Paas, F. G. W. C. (2002). Redirecting learners' attention during training: Effects on cognitive load, transfer test performance and training efficiency. *Learning and Instruction, 12*(1), 11–37.
van Merriënboer, J. J. G., Kirschner, P. A., & Kester, L. (2003). Taking the load off a learner's mind: Instructional design for complex learning. *Educational Psychologist, 38*(1), 5–13.
Vaughn, K. E., & Kornell, N. (2019). How to activate students' natural desire to test themselves. *Cognitive Research: Principles & Implications, 4*, 35.
Velan, G. M., Jones, P., McNeil, H. P., & Kumar, R. K. (2008). Integrated online formative assessments in the biomedical sciences for medical students: Benefits for learning. *BMC Medical Education, 8*(1), 52.
Wahlheim, C. N., Finn, B., & Jacoby, L. L. (2012). Metacognitive judgments of repetition and variability effects in natural concept learning: Evidence for variability neglect. *Memory & Cognition, 40*(5), 703–716.
Wissman, K. T., Zamary, A., & Rawson, K. A. (2018). When does practice testing promote transfer on deductive reasoning tasks? *Journal of Applied Research in Memory and Cognition, 7*(3), 398–411.
Yan, V. X., Bjork, E. L., & Bjork, R. A. (2016). On the difficult of mending metacognitive illusions: A priori theories, fluency effects, and misattributions of the interleaving benefit. *Journal of Experimental Psychology: General, 145*(7), 918–933.
Yeo, D. J., & Fazio, L. K. (2019). The optimal learning strategy depends on learning goals and processes: Retrieval practice versus worked examples. *Journal of Educational Psychology, 111*(1), 73–90.

The order of participants' choices indicated that they sometimes used the restudy opportunities as feedback following the test opportunities, particularly for definitions they could not recall.

Similar interactions have been reported in memory-based tasks, where learners with high prior knowledge show larger benefits of retrieval over restudy (Carpenter et al. 2016), and learners with higher fluid intelligence show larger benefits of retrieval for difficult items compared to easier items whereas learners with lower fluid intelligence show larger benefits of retrieval for easier items than for difficult items (Minear et al. 2018).

Affiliations

Shana K. Carpenter[1] **· Tino Endres**[2] **· Luotong Hui**[3]

[1] Department of Psychology, Iowa State University, W112 Lagomarcino Hall, 901 Stange Road, Ames, IA 50011, USA

[2] Department of Educational and Developmental Psychology, University of Freiburg, Freiburg, Germany

[3] Department of Educational Development and Research, Maastricht University, Maastricht, Netherlands