Do students space their course study? Those who do earn higher grades.

Marissa K. Hartwig[1] and Eric D. Malain[2]

[1] University of South Florida
[2] Washington State University

Learning and Instruction

In press version 2021

Corresponding author:
Marissa K. Hartwig
Department of Psychology
University of South Florida
Email: mhartwig@usf.edu

Eric D. Malain
Department of Psychology
Washington State University

Funding:
This research was supported by the Institute of Education Sciences, U.S. Department of Education, grant R305A130699 to Washington State University and grant R305A160263 to the University of South Florida. The opinions expressed are those of the authors and do not represent views of the Institute or the U.S. Department of Education.

Acknowledgments:
Thank you to Joyce Ehrlinger for helping to make this research possible; to Pepper Williams for facilitating access to data; and to John Dunlosky for comments on an earlier draft of this manuscript.

Declarations of interest:
None.

Abstract

We examined students' naturalistic decisions about spacing their study in an undergraduate course (N=185) and whether self-selected spacing predicted course performance. Usage of two study tools – an online textbook and quiz tool – was recorded daily. We operationalized spacing as how often the tools were used and the timing of their use relative to exams. We found that students increased their study near deadlines and exams, used the textbook more often than the quiz tool, and used the tools infrequently when they were optional (vs. required). Importantly, spaced retrieval practice (via quiz tool) predicted course performance and GPA, whereas spaced reading (via textbook) was a weaker predictor. That is, when students opted for more frequent and early quizzing, they earned higher grades, even controlling for time spent quizzing. Thus, self-selected spaced study – especially spaced retrieval practice – supports student achievement.

Keywords: spaced study, self-selected spacing, college students, course performance

Do students space their course study? Those who do earn higher grades.

1. Introduction

Spacing of study refers to spreading, or distributing, one's study of the same materials across time. For a given amount of time spent studying the same materials, study can be spread into many sessions (more spacing) or packed into only one or few sessions (less spacing). Figure 1 illustrates this idea and also shows that a given number of study sessions can be spread across a longer span of time (more spacing) or across a shorter span of time (less spacing). Decades of research on spacing have shown that studying with a greater degree of spacing produces better long-term retention compared to studying that is massed or closely spaced together (for reviews, see Cepeda et al., 2006; Dempster, 1989). Importantly, spacing has been investigated with experimental designs in which participants are randomly assigned to study conditions of more or less spacing, while holding constant the total study time and the test delay (i.e., the time from the last instance of study until the test). In these experiments, higher test scores can be attributed to more spacing rather than to differing amounts of study or differing test delays. This benefit of spacing, known as the spacing effect, has been demonstrated for a variety of tasks and materials – including verbal information (e.g., Bahrick, 1979; Sobel et al., 2011), mathematics (e.g., Hopkins et al., 2016), and motor tasks (e.g., Moulton et al., 2006) – and across a wide range of procedures and learner characteristics (see Dunlosky et al., 2013). Most research on spacing has been conducted in lab settings, but spacing effects have been shown in classrooms as well (e.g., Carpenter et al., 2009; Hopkins et al., 2016; Seabrook et al., 2005). Thus, researchers have rightly advised students and teachers that long-term learning can be enhanced with spaced study, even when total study time remains unchanged.

However, less is known about students' implementation of spaced study when they select it for themselves in an educational setting, like when studying in a college course. Much research in educational settings has explored students' choices about allocating time to their studies (Romero & Ventura, 2020), but this research has not specifically examined spacing. Some classroom-based studies have experimentally manipulated spacing and test delays, for a given amount of study, but they did not investigate students' own choices with regard to spacing. While spacing manipulations are essential for demonstrating cause-and-effect, they impose behaviors that may not be natural for students. Indeed, some students may have difficulty following a spaced schedule when they must implement it themselves (Barzagar Nazari & Ebersbach, 2018).

Furthermore, based on prior research, it is unclear whether students' self-selected spacing would be a good predictor of grades in a real-world college course. On one hand, as described above, a large body of experimental research shows that greater spacing corresponds to better learning outcomes (e.g., Cepeda et al., 2006). But, on the other hand, countless factors can affect student outcomes, especially in real-world uncontrolled educational settings, and thus the advantage of spacing might be minimal or simply undetectable against this backdrop of influences (e.g., differences in student ability, prior knowledge, motivation, competing life demands, and much more). Furthermore, spacing effects are most apparent for tests that occur after a substantial test delay, and, in fact, spacing effects can be reduced, eliminated, or sometimes reversed with very short test delays (Rohrer, 2015; Cepeda et al., 2006). In most college classrooms, students can study until almost immediately before an exam occurs, so spaced study may not benefit performance on those exams due to the extremely short test delay, perhaps enabling cramming (massing) to be a successful strategy. Thus, for multiple reasons, self-selected spacing in a college course is not guaranteed to predict students' grades.

The present study focused on students' self-selected degree of spaced study and whether it was associated with grades. Our first goal was descriptive: We employed behavioral measurement to paint a detailed picture of the extent to which students choose to space their study in a college course. Our second goal was to examine whether self-selected spacing would predict grades. Before presenting our study, we first review what is known about students' choices to space their study and under what circumstances, and we then consider in more detail the possible link between self-selected spaced study and grades.

1.1. Do students choose to space their study?

1.1.1. Surveys and self-reports. Self-reports from college students indicate that cramming before an exam or deadline is common practice. In recent surveys, 53% (Morehead et al., 2016) and 66% (Hartwig & Dunlosky, 2012) of undergraduates reported that they regularly cram large amounts of information the night before a test. In another survey, undergraduates in introductory courses rated how much they used a variety of study strategies, and spaced practice was reportedly used less often than other strategies like practice testing (Bartoszewski & Gurung, 2015; see also Susser & McCabe, 2013). Nonetheless, many students do believe spaced study is beneficial for learning (Blasiman et al., 2017; Cohen et al., 2013, experiment 7; Taraban et al., 1999). For example, in a survey of college students by Susser and McCabe (2013), approximately 85% said that spacing study (rather than massing the same amount of study into one session) is better for long-term retention of materials. When asked about reasons why they either cram or space their study, students indicated they space their study more when the upcoming test is difficult, when the test is heavily weighted in their course grade, when the test material has high value for future courses or careers, and when there is a large amount of material to be learned. Spacing of study was also associated with an individual's use of elaborative study strategies and higher levels of metacognitive self-regulation. Thus, although students may have certain tendencies (e.g., cramming for low-stakes tests), individuals differ in the extent to which they choose to space their study.

Self-reports are valuable indicators of how students perceive their study and likely have some association with actual behavior, but self-reports are vulnerable to inaccuracies (Nisbett & Wilson, 1977). Some studies have used daily diaries to provide a more detailed and accurate self-report, and although these studies did not focus on spacing per se, frequency and timing of study have been found to be associated with GPA (Nandagopal & Ericsson, 2012). However, when students report their study, it is unclear to what extent the total amount of study time varies along with spacing, an issue which we discuss again below.

1.1.2. Laboratory research. Lab studies, also, have investigated students' choices about spacing their study (e.g., Ariel et al., 2009; Benjamin & Bird, 2006; Cohen et al., 2013; Pyc & Dunlosky, 2010; Son, 2004; Toppino & Cohen, 2010). However, these lab studies are typically one-session experiments using simple learning materials (e.g., word pairs) and may have limited applicability to college students' decisions about spacing complex materials across days or weeks. Nonetheless, we give a brief overview here. In these studies, participants briefly study to-be-learned items and then choose whether to study an item again immediately (mass) or later in the same session (space), prior to a final test of cued recall. Researchers have varied the duration of item presentation, item value and difficulty, and constraints on choice. Under most conditions, spacing was preferred over massing, but participants may prefer spacing because they do not want to see the same item again immediately or because they want to restudy items closer to testing (for discussion, see Cohen et al., 2013; Toppino & Cohen, 2010). Indeed, when test delay was held constant for each item, participants preferred shorter rather than longer spacing (Cohen et al., 2013). While laboratory studies can provide precise experimental control for investigating spacing, including the ability to control the total amount of study, participants' choices in these studies are not necessarily equivalent to choices made in real-world educational settings, where studying may occur across days or weeks and students have many motivations and demands on their time.

1.1.3. Behavioral measurement in educational settings. Numerous studies have gathered behavioral trace data to examine students' allocation of time in educational settings. In many of these studies, computer-based learning environments or online tools have been used to collect students' moment-to-moment choices regarding their study strategies, plans, and monitoring of their learning and have provided valuable information about students' self-regulated learning (for an overview, see Greene

& Azevedo, 2010). For example, Taraban et al. (1999; studies 2 and 3) gathered behavioral data from undergraduates in psychology courses that used online materials including self-paced lessons, interactive activities, quizzes, assignments, and webpage resources. They measured – on a daily basis across an entire semester – the time that students spent using the online materials and the frequency of their electronic activity. Students' behavior revealed that the online resources were used most heavily during the two days preceding exams. Also, when interim deadlines for quizzes and assignments occurred between exams, students' activity was concentrated immediately before each deadline (also called a "scallop" pattern; see Michael, 1991), indicative of cramming. The researchers did not report whether variance in this behavior predicted student performance.

In other studies, trace data in online or hybrid classes have been used to examine how students' study choices affect their course performance. For example, learning analytics and educational data mining are both active fields of research in which studies aggregate information from student interactions with technology, often to create models that predict student success or failure (Romero & Ventura, 2020). Using these methods, researchers have examined relationships between student engagement with course content and course outcomes. They have found that consistently engaging with course content throughout a semester, compared to low activity, predicts better course outcomes (van Leeuwen et al., 2019). Similarly, regular or consistent engagement with pre-class activities in flipped courses predicts positive course outcomes (Jovanovic et al., 2019). Studies like these typically include many indicators of student engagement, which are entered into regression models to identify activities that predict course outcomes, and efficacy of specific activities can vary widely across different courses (Gašević et al., 2016). These studies commonly include quizzing among their activities, but the quizzes are typically used as an assessment of student knowledge on a topic (Romero & Ventura, 2020) rather than a strategy for learning (see Karpicke & Roediger, 2008). Most important, we are unaware of any trace data studies that have examined the impact of spacing per se.

1.2. Is spaced study associated with higher grades?

Prior studies have not examined whether self-selected spacing predicts grades. As mentioned above, the impact of spaced study on grades might be minimal when test delays are short or when student performance is affected by many real-world factors. Still, there are plausible reasons that spacing might predict grades, as we describe next.

1.2.1. Theories of spacing effects. Countless experimental studies have demonstrated an advantage of spaced practice (vs. practice that is massed or less spaced) for performance on delayed tests, and numerous theories have been proposed to explain these effects. The theories have featured ideas such as encoding variability, consolidation, deficient processing, study-phase retrieval, and more. A review of possible mechanisms is outside the scope of this paper (for detailed reviews, see Benjamin & Tullis, 2010; Delaney et al., 2010), and we do not aim to evaluate these mechanisms. For illustrative purposes, however, consider how study-phase retrieval might help to explain why spaced study enhances student learning. According to this explanation, students retrieve, from memory, elements of previous study sessions during subsequent study sessions, and this act of retrieval strengthens memory for the restudied materials – especially when retrieval is effortful. Thus, when study is spaced into more sessions or across longer spans of time, students have the chance to forget information in the intervals between study sessions, and subsequent study sessions produce the effortful retrieval that enhances encoding. Interestingly, such a mechanism might also imply that study activities that more strongly encourage retrieval would magnify the benefit of spacing compared to more passive study activities. Regardless of the exact mechanism, however, it is clear that spacing effects are robust and can impact student learning.

1.2.2. Self-selected spacing. When spacing is self-selected by students, third variables may also contribute to an association between spaced practice and performance. For example, students who are typically high performers may be more likely to space their study, perhaps because they have higher motivation or better self-regulatory skills, compared to low performers. Also, greater spacing of study may allow use of other strategies not allowed by cramming (e.g., asking the instructor questions, getting help from peers, having time to ponder complex ideas, getting adequate sleep) which may further support student learning. These potential downstream advantages of spaced study are possible in real-world contexts and might have interesting educational implications. Thus, when examining a possible association between self-selected spacing and performance, ruling out all third-variable explanations is not our goal. Demonstrating the association is an essential starting point for future research into potential explanatory variables. To better understand the association, it is necessary to control for one particular variable – time spent studying – because a student who starts studying sooner before an exam or studies on more separate days (i.e., behaviors that signal spaced study) might also study more time overall compared to another student who waits to begin studying until the day before an exam. By controlling for total study time, any relationship between spacing behaviors and performance would indeed reflect spaced study rather than merely a greater investment of time.

1.2.3. Spacing of different study activities. Spacing refers to a distributing of restudy, but it does not specify a particular activity that must occur during a restudy session. It is plausible that the strength of spaced study as a predictor of student performance may depend on the type of study activity. One form of study frequently self-reported by students is to reread materials (Karpicke et al., 2009; but see Kuhbandner & Emmerdinger, 2019), but rereading is a relatively poor study activity. An abundance of research on the testing effect (for reviews, see Adesope et al., 2017; Dunlosky et al., 2013; Roediger & Karpicke, 2006) has demonstrated that materials are better retained when restudy sessions involve attempts to retrieve information from memory (e.g., practice tests) rather than merely rereading. Thus, spacing of retrieval practice, a highly efficacious study activity, may predict learning outcomes to a greater extent than spacing of rereading.

1.3. Overview of the present study

The present study had two main questions: (1) To what degree do students choose to space their study in an authentic educational setting such as a college course? (2) Do differences in self-selected spacing predict student performance? To address these questions, we measured a subset of students' study behaviors – i.e., their use of an online textbook and their use of an online quiz tool – in a large, undergraduate social psychology course. These two study tools were provided to all students in the course, and usage was measured daily throughout an entire semester. These tools contrast with the online materials and tutorials used in some other trace data studies in that they mimic conventional materials relied on heavily by students in traditional in-person courses (e.g., textbook, study guides, practice questions). By measuring students' use of these online tools across an entire semester, we gained access to a large portion of their naturalistic, self-selected spacing of study in a college course.

To address our first question, we examined students' usage of the tools across the semester to reveal how early they accessed materials and how often they returned to those same materials, which provided evidence about spacing behaviors. We expected to find that college students would engage in a high degree of cramming and would be strongly influenced by deadlines, yet we also expected to see variability in students' spacing behavior. This variability in behavior led to our second question, which we addressed by examining the association between spacing and student academic performance. When assessing the association, we controlled for total time spent using the study tools. Also, we examined the two study tools separately because spacing of a highly effective study behavior (such as retrieval practice) might be more beneficial than spacing of a less effective study behavior (such as rereading). Our main hypothesis was that greater spacing would be associated with higher student performance – for reasons that could include spacing effects as well as associations with unmeasured third variables.

2. Method

2.1. Participants

Participants were 191 undergraduate students (60.7% female, 8 unreported) enrolled in a social psychology course at Washington State University. The course was an upper-division undergraduate psychology course for social science majors. Almost half of the class (43.5%) were juniors, and the remainder was split between sophomores (24.1%) and seniors (28.3%) (8 unreported). The mean age was 20.7 years (*SD* = 1.5). Participants were 58.1% white, 7.3% Black, 5.2% Asian, and 8.9% other (20.4% unreported); they were also 21.5% Hispanic (9 unreported). Six students were eliminated from analysis because they did not complete the course in its entirety, leaving a final sample size of 185 students.

2.2. Procedure

At the start of the semester, two tools – an online textbook and online quiz tool – were provided to the entire class free of charge by the publisher, W. W. Norton & Company, in return for the instructor agreeing to test the beta version of the quiz tool in this course. The textbook was the online version of Social Psychology 4th Edition (Gilovich, Keltner, Chen, & Nisbett, 2015), and the quiz tool platform (called *InQuizitive*) was tailored for this text. The online textbook (henceforth called e-book) and quiz tool could be used at any time of students' choosing. Starting during the second week of the semester (January 19th), we downloaded a record of student usage of both the e-book[1] and quiz tool each day just after midnight. Thus, we could determine whether each student had used either tool during the 24 hours between each download (i.e., one-day intervals).

Five exams were administered during the semester, and each exam covered 2-3 chapters. For all 11 chapters, e-book readings and quiz tool modules were available for student use. For the first chapter relevant to each exam (i.e., Chapters 1, 3, 5, 6, & 7), students were awarded points for using quiz tool modules; we will refer to these chapters as *required*. The remaining chapters (i.e., Chapters 2, 4, 8, 9, 10, & 11) also had e-book readings and quiz tool modules available, but students did not receive credit for using them; we refer to these chapters as *optional*. Note, however, that all of these chapters were tested on exams, students understood they were expected to master all of them, and the instructor encouraged students to use the e-book and quiz tool for all chapters.

Figure 2 shows student usage of tools each day of the semester, denoting exam dates and assigned deadlines for completing required quiz tool modules. Exams were spaced about three weeks apart on average. We tracked usage of each tool for the chapters relevant to the upcoming exam. For

[1] While it is possible students could have purchased a print copy, it is unlikely given the high cost of print copies and free access to the online textbook.

example, beginning the day after Exam 1 and continuing through the day of Exam 2, we tracked tool usage for chapters 3 (required) and 4 (optional). Any studying of those chapters that occurred prior to the relevant time period, which was rare, is represented in the first recorded day of the time period (i.e., January 28th for the second exam period).

Immediately following exams, students completed post-exam questionnaires, which were attached to the end of the printed exam (see Materials below). Students were also surveyed during the fourth week and eighth week of the semester to collect self-reported GPA, demographics, and other information unrelated to this study.

2.3. Materials

The materials for this study included two online study tools, five in-class exams, five post-exam questionnaires, and two online surveys.

2.3.1. Online study tools. All students were given access to two study tools: an online textbook (e-book) and an online quiz tool. E-book material and quiz tool material largely overlapped with lecture material. The course syllabus provided a lecture schedule with assigned e-book readings (i.e., subsections of chapters) and recommended that students read these sections before lecture. Students could access the e-book via the internet at any time, and completion of this task was not rewarded with points or known to the instructor.

Students could also access the quiz tool via the internet at any time and could use it as often as they wanted. The content of the quiz tool modules was designed to closely correspond to the content of assigned reading subsections of each chapter. This content was a subset of the total lecture material covered in class and represented on exams. The quiz tool had a bank of 13-34 quiz questions per chapter, but because not all subsections of a chapter were assigned, students in the present study had access to approximately 10-18 of these questions per chapter. Therefore, students likely saw repeated items. For required chapters, students had to earn a requisite amount of points via the quiz tool for full credit.

The quiz tool employed multiple-choice, matching, and fill-in-the-blank questions. Correct responses were marked by a green indicator and a dialogue box with a restatement of the correct answer concept. Incorrect responses were marked by a red indicator and students were prompted to try again until the correct answer was identified. Question selection was fully controlled by the tool, which favored selection of target concepts that students previously answered incorrectly. Student control was limited to when or whether to use the quiz tool.

2.3.2. Exams. Five non-cumulative exams were administered during the semester. Exams were closed-book and completed during the class period (lasting approximately 75 minutes). Each included 21 multiple-choice questions worth one point each and three short-answer questions worth three points each, totaling 30 points. Like many college exams, question types included knowledge questions (recalling factual information), conceptual questions (understanding of individual concepts), application questions (applying concepts to novel situations), and analysis questions (breaking down concepts into constituent parts and identifying relationships among those parts). Each exam had acceptable internal consistency (Cronbach's alpha of .70 - .76). The mean percent scores for exams 1 through 5 were 76.0% (*SD*=14.3%), 79.5% (*SD*=15.4%), 74.2% (*SD*=15.1%), 72.0% (*SD*=14.2%), and 77.4% (*SD*=15.6%), respectively. Each student took four or five exams. Students had the option to miss one exam during the semester[2] and take a more difficult, free-response, make-up exam during the end-of-semester finals period. Due to nonequivalence, make-up exams are not included in our measure of course performance; we defined a student's course performance as their average score across regular exams only (Cronbach's alpha = .86).

2.3.3. Post-exam questionnaires. A one-page questionnaire was attached to each exam and was completed immediately following the exam. The questionnaires included several brief questions concerning the student's exam preparation and estimated performance. Most relevant here, on the questionnaires following exams 4 and 5, we asked students to estimate their total time spent studying outside of class, as well as the percentage of that time spent on various study activities – including the quiz tool and e-book (see Appendix A). Students indicated that their total study time for this course (in any form outside of lecture) since the previous exam was a mere 5.7 hours on average. Importantly, time spent with the e-book and quiz tool was a substantial fraction of students' self-reported study: 32% of their time was spent reading or studying the e-book; 21% was spent using the quiz tool; 41% was spent studying their lecture notes; and the remainder was spent on other strategies or activities. Thus, according to these self-reports, the combination of e-book and quiz tool usage constituted more than half (53%) of students' total time spent studying.

2.3.4. Online surveys. Additional information was collected via two online Qualtrics surveys. The first survey (distributed February 2nd) included demographic questions, as well as other items not relevant to the present study (e.g., nicknames; psychological scales for another project). The second

[2] Out of 185 students, 4 missed Exam 1 (2.2%); 10 missed Exam 2 (5.4%); 5 missed Exam 3 (2.7%); 23 missed Exam 4 (12.4%); and 9 missed Exam 5 (4.9%).

survey (distributed March 4^{th}) asked students to report their college GPA and also included scales not relevant here.

2.4. Operationalizing spacing

When examining students' usage of the two online study tools, we looked for evidence of spacing in two ways that correspond to the two panels of Figure 1 showing how students might space their study: First, we sought to capture the degree to which students spaced their study of the same materials into separate study sessions (as in Figure 1, top panel). Thus, we computed the number of days on which each student used the tools for a given chapter, which we will call *frequency* of tool use. For example, a frequency of 2 for the chapter 5 quiz tool means that the student used that tool twice during the multi-week period before the exam. Though other studying may have occurred outside of the tool, this frequency nonetheless represents observable evidence of spaced study[3].

Second, we sought to capture the span of time across which students restudied chapter materials – starting with the earliest evidence of tool use for a given chapter and concluding with the last instance of study before the exam (as in Figure 1, bottom panel). Typically, the last instance of study for nearly all students occurs during the 24 hours preceding an exam, because most students review to-be-tested concepts shortly before an exam (see Taraban et al., 1999). However, this last instance of study is often not reflected in our record of tool usage, because the study tools provided may not have been all students' preferred method of exam review (compared to, say, reviewing lecture notes). Thus, when estimating span, we relied on the key assumption that the last instance of study occurred shortly before the exam for all students (i.e., it was a constant), and therefore span was solely determined by the earliest evidence of tool use. Thus, we measured the number of days before the exam that a student first used the study tool – a variable we will call *first use*. For example, a first use of 10 for the chapter 5 quiz tool means that the student first used the tool 10 days before the exam. We believe that first use is a reasonable approximation of a student's span of study. (However, for readers who reject the assumption that the last instance of study occurs shortly before an exam, the variable of first use would instead simply represent how early a student began studying. At the least, earlier first use would signify greater *potential* for spacing to occur.)

While no operationalization of spacing is perfect in a naturalistic and uncontrolled setting, these two variables (i.e., frequency and first use) represent an attempt to obtain behavioral evidence of two distinct aspects of spaced study (as in Figure 1). Of course, because students' behavior was naturalistic,

[3] To the extent that other (unobserved) studying occurred outside of tool use, our ability to detect a relationship between our spacing variables and student performance would be reduced.

it is likely that their total study time also varied. If we had mandated a particular amount of study time, this requirement would undermine the degree to which study was naturalistic. To account for this necessary lack of experimental control, we employed a statistical control of time when examining whether spacing predicted student performance. That is, we controlled for the amount of time (in minutes) a student used the tool.

Also, we restricted analyses to required chapters only (no optional) when examining the relationship between spacing and student performance. We did this because student behavior differed markedly between required and optional chapters, as shown in the Results (see also Figure 2). In fact, with optional chapters, tools were often completely unused, which would either extensively reduce our sample (if we limited analysis to users only) or would make tool use vs. disuse a main source of variability in our measures (rather than spacing). For required chapters, however, most students did indeed use the tools, and variance in our measures reflected the degree of spaced practice we sought to understand. In other words, by focusing on required chapters, we were able to ask: Among users of a tool, did greater spacing predict exam performance?

We computed an average for each behavior described above (i.e., frequency, first use), representing a student's spacing behavior across the entire semester. We believe that averaging across the semester is more appropriate than repeating analyses exam-by-exam for several reasons: As noted previously, the exam periods differed considerably from one to the next, including differences in the topics covered, difficulty of materials, length of exam periods, timing of deadlines, and segment of the semester (e.g., early in the semester vs. near midterms or finals). Thus, individual exam periods are not well-suited for comparison against each other. Furthermore, these differences introduced noise in the spacing and performance measures, making associations more difficult to detect on an exam-by-exam basis. Instead, semester averages are better suited to capture general and reliable patterns of behavior and performance across time.

Thus, the following two composite variables were computed for each student, separately for each tool, across required chapters only: *Frequency* was the number of days on which a student used the tool in an exam period, on average. *First use* was how long before an exam (in days) a student first used the tool, on average. These two composite variables were moderately correlated (r = .43, p < .01 for the quiz tool; r = .69, p < .01 for the e-book), yet they appear to capture differing aspects of student choice about spacing. That is, frequency captures the spread of study across more sessions, whereas first use aims to capture the spread of study across a longer span of time, as in Figure 1. These two composite variables were examined as predictors of student performance. We had two measures of student performance: *Course performance* was a student's average exam score across the semester, and *college GPA* was a student's self-reported college GPA from an online survey. Finally, we also computed *tool time*[4] – a student's behaviorally-measured total time (in minutes) spent using each tool on average during an exam period – which we used as a control in our analysis of the relationship between spacing and student performance.

3. Results

3.1. To what degree did students choose to space their course study?

We describe students' spacing with respect to the two online study tools (i.e., e-book and quiz tool), showcasing *how often* (frequency) and *how early* (first use) the tools were used.

3.1.1. How often. Table 1 shows how often students used each tool, organized by chapter. For chapters with a required quiz-tool assignment, nearly all students used the quiz tool (96%) and most also used the e-book (82.4%). Among the users, most students used the quiz tool on only one day, and the e-book was commonly used on one to three days. In other words, the e-book was typically used more times than the quiz tool. When assignments were optional, however, the number of users decreased dramatically for both the quiz tool (31.8%) and e-book (54.6%). Among users, the most common frequency was a single day of use. Thus, most students did not space their use of each tool across many days. Furthermore, students may be disinclined to use quiz tools or read chapters unless an instructor creates an external incentive (e.g., points) for doing so. When usage was optional (i.e., no points), many students did not take advantage of the tools provided to support their learning – despite needing to learn all eleven chapters.

Furthermore, an interesting question arises for the required chapters: Did students ever use the tools after the deadline had passed in preparation for the exam? Among students who used the quiz tool, few used it post-deadline (18.6%); among students who used the e-book, however, many used it post-deadline (71.2%) – and with higher frequency than the quiz tool. Thus, students were more likely to return to the e-book beyond the deadline as they prepared for an exam.

3.1.2. How early. Figure 3 shows how long before an exam (in days) a chapter tool was first used. First use differed considerably by chapter, and this was likely due to the course schedule (refer to Figure 2). Thus, we highlight only general trends without dwelling on specific chapters. Tool use for required chapters generally began earlier than for optional chapters. We suspect this occurred for at least two reasons: (1) Required chapters corresponded to lectures that occurred earlier in each exam

[4] Descriptive statistics for tool time can be found in Table 2 and will be discussed more later.

period. (2) The instructor gave a deadline for earning credit on the required quiz-tool assignment, and this deadline typically occurred a week or more before an exam. In sum, how early students used the tools appeared to fit the demands of the course.

3.2. Did self-selected spacing predict student performance?

Table 2 shows descriptive statistics for the composite variables representing spacing, as well as the relationships between these variables and our two measures of student performance. Regarding the quiz tool, both spacing behaviors (i.e., greater frequency and earlier first use) were associated with higher student performance. Importantly, these associations remained significant when controlling for the total time spent using the tool and, therefore, cannot simply be attributed to overall larger amounts of study. In fact, time on the quiz tool was unrelated to performance in bivariate correlations (Table 2). Furthermore, when used as a control, time on the quiz tool was *negatively* related to performance (discussed below).

Compared to the quiz tool, the total time spent using the e-book was about three times longer (102 min vs. 34 min), consistent with prior research indicating that students spend more time reading than quizzing themselves (e.g., Karpicke et al., 2009). Despite spending more time, e-book spacing was less predictive: First use was weakly related to performance and frequency was unrelated (Table 2). Indeed, reading chapters earlier, perhaps before class, could plausibly benefit students' comprehension during lecture. Nonetheless, e-book spacing was a weaker predictor of performance than quiz tool spacing[5].

4. Discussion

In this study, we used behavioral measurement to investigate students' self-selected spacing of their study in a college course across an entire semester. By measuring behavior daily, we were able to obtain a high level of descriptive detail without burdening students with daily self-reporting or relying on the accuracy of their retrospective memory. Our results complement conclusions from previous research. We found that students underutilized spacing by spreading their study over relatively few days. Moreover, many students failed to use non-incentivized (optional) study tools, whereas nearly all students used required tools at least once. Importantly, students' self-selected spacing predicted their course grades and college grades. For instance, students earned higher grades if they spaced their quizzing more – that is, quizzed earlier and more often, even controlling for time. Thus, the present

[5] For interested readers, exam-by-exam analyses are presented in Appendix B. Though weaker in magnitude, the exam-by-exam associations are generally consistent with the composite analyses presented here in the Results.

research highlights the importance of spaced study, not merely a greater investment of time, for predicting grades in an authentic educational context in which students have control over their study.

Spaced study was a stronger predictor of grades when the study activity was retrieval practice (quizzing) rather than reading. Compared to quizzing, students spent more time reading the e-book, read it on more separate days, and showed more variability in reading behavior. Nonetheless, students' spacing of their reading was a weaker predictor of student performance. Of course, the correlational design of our study does not permit a statistical comparison of the two study activities, but the predictive strength of spaced retrieval practice is consistent with both intuition and previous findings. Indeed, it is intuitive that spacing an effective study strategy would be more potent than spacing a less effective strategy, and an extensive literature has shown retrieval practice to be a more effective strategy than rereading (i.e., testing effect; Roediger & Karpicke, 2006). Note that the observed pattern also fits with the possibility of a study-phase retrieval mechanism of spacing, which relies on activating memory traces from prior study sessions. Such activation may occur to a greater extent with testing or quizzing. Relatedly, Rawson, Vaughn, and Carpenter (2015) showed that spaced practice and retrieval practice can interact. Specifically, they found that the benefit of retrieval practice (vs. rereading) is larger when practice is more spaced (vs. less spaced). Indeed, spaced retrieval practice promises to be a potent combination of learning techniques.

Nonetheless, many students quizzed themselves only one day per exam period, often just before the deadline – displaying little spacing. Of course, there is no absence of spacing in the sense that the average student was exposed to the same material multiple times (e.g., reading, lecture, quizzing, reviewing notes – perhaps totaling 4-6 exposures). Still, students underutilized the opportunity to space their quizzing. That is, they could have reached higher levels of spacing by quizzing across more days. The choice to quiz only once resonates with lab findings in which many participants drop items from restudy after only one successful retrieval, even though repeated retrieval yields better testing outcomes (Dunlosky & Rawson, 2015). In other words, students make suboptimal study choices in both the lab and classroom. These suboptimal choices manifested here in single use of required tools and complete abstinence from optional tools for many students.

What are the reasons for these poor study choices? One possible reason is that students may not know what study behaviors are most effective for learning. For example, some students may not recognize the value of spaced retrieval for supporting memory of to-be-learned materials. In fact, previous research suggests that students undervalue spaced practice (e.g., Cohen et al., 2013) and often view testing simply as a diagnostic tool rather than as a valuable learning technique (e.g., Kornell &

Bjork, 2007). Also, previous research suggests that students may perceive challenging study activities (such as spaced retrieval) to be less effective than activities that feel easy. In other words, they do not recognize that difficulties can be *desirable difficulties* that enhance learning (Bjork, 1994). A second possible reason for poor study behaviors pertains to one's ability to implement a desired strategy. Indeed, students are often overscheduled with heavy course loads, extracurricular activities, and jobs – which may pose challenges for spacing their study. Also, they may experience costs associated with task switching if study is heavily spaced. Further, even if a student wants to space their study into more sessions or begin studying sooner, they might fail to space their study if their time management is poor. Indeed, time management has been linked to student success, and behaviors such as planning and scheduling (which could support spaced study) are commonly included in measures of time management (e.g., Britton & Tesser, 1991; Macan et al., 1990). If students with poor time management skills struggle to space their study, a potential route to encourage more spaced study might involve training to develop self-regulatory skills that support time management (e.g., Jansen et al., 2020). Although speculative, interventions that emphasize planning, as well as building knowledge of effective study techniques like spacing, might help to increase spaced practice when students control their own study. Of course, spacing can also be encouraged by imposing class requirements for restudy (see Implications).

Because the present study investigated naturalistic, self-selected study behavior – meaning that students chose how and when they studied – we must consider study time when examining how spacing relates to performance. Students themselves determined how long they used each study tool, and we measured this amount of time and used it as a statistical control. Many previous studies of spaced practice, especially those in laboratory settings, have experimentally controlled how long a participant studies (requiring equal time across conditions) to ensure that the amount of time spent studying is not the underlying cause of different student outcomes. Indeed, one might reasonably expect that more time studying would be associated with better outcomes. However, this intuition may not always be correct when students control their own study time (see Plant et al., 2005). In the current study, for instance, there was no bivariate relationship between study time and performance outcomes. Interestingly, when time spent quizzing was used as a control, time was negatively related to performance. There are many possible explanations for a negative relationship between time and performance. For example, it might be explained by engagement with retrieval practice, because attempting to access memory during a quiz might take less time than looking up every answer but may produce better retention. Furthermore, students who struggled with the to-be-learned concepts likely took longer to earn the point requirement. In other words, when study is self-regulated, more time spent quizzing might sometimes reflect inefficient or ineffective study.

4.1. Limitations and future directions

In the present study, students' study behaviors were self-selected rather than experimentally assigned, and while this is a strength for describing students' real-world choices, it limits our ability to explain why spaced study and student performance were associated. The association may partly reflect well-established spacing effects, but it may also reflect the effects of unmeasured third variables. For example, how much an individual spaces his or her study may be associated with student characteristics (e.g., motivation) or external factors (e.g., light course load) that also help performance. One intriguing possibility is that spaced practice may be associated with engaging in other beneficial behaviors or strategies (e.g., help-seeking, planning, pondering ideas) that would be less feasible when students cram. To the extent that spaced practice enables such behaviors in the real world, this may be an important but underappreciated benefit of spaced practice – and a fertile ground for future research.

Since we had access to students' online behaviors only, our measure of study behavior captured only a portion of students' studying. While the two study tools we measured constituted more than half of students' study activities according to self-reports, other unmeasured activities or student characteristics could also affect performance. For example, we did not measure students' prior knowledge or their success on practice questions, both of which may predict exam performance. However, it is unclear how or if these unmeasured variables covary with spacing choices. One possibility is that more prior knowledge would predict more spacing (e.g., more frequent use of quizzing), perhaps because quizzing allows knowledgeable students to quickly evaluate their understanding and provides positive feedback for knowledgeable students. If so, prior knowledge might offer an alternative explanation for the link between spacing and grades. Then again, previous research suggests that students are unlikely to repeatedly test themselves on materials they have already answered correctly once (e.g., Dunlosky & Rawson, 2015). Thus, another possibility is that more prior knowledge or success on practice questions might encourage less restudy (because students perceive their learning to be sufficient). If so, prior knowledge would be associated with less spacing yet still high course performance, thereby working against a positive association between spacing and grades.

Also, it is unclear how relatively important the online study tools were for students' overall study. If these online activities were less important than other study activities, this would work against finding a relationship between spacing of these activities and performance. Even so, the activities we measured were predictive of student performance. Furthermore, students who space their use of study tools may be likely to space other (unmeasured) study activity as well. Plausibly, the strength of association between spacing and performance may have been stronger if we had been able to observe all study behavior.

The behavioral evidence available in the present study – i.e., students' use of the e-book and quiz tool – allowed only an imperfect operationalization of spacing. First, some rereading and retrieval practice may have occurred outside of the online tools. For example, students might reread their lecture notes on the same course material as the e-book readings. Similarly, students could quiz themselves with flashcards or some other self-generated tactic. Indeed, our operationalization of span (i.e., first use) assumes that students restudy materials outside of the online tools shortly before an exam. Second, we do not know exactly what students did while using the online tools. For example, a student might not have been reading the entire time the e-book tool was open. Also, using the e-book multiple times does not guarantee rereading because students could read different segments of a chapter each time they used the tool. Retrieval practice is also not guaranteed by using the quiz tool because students might look up answers instead of attempting retrieval. Each of these possibilities reduces the likelihood that we would find evidence of the benefit of spacing.

The structure of a course will inevitably affect students' behavior. In the course observed here, for instance, the deadlines for practice quizzing encouraged students to engage in some spaced practice that otherwise might not have occurred. Similarly, spacing behavior may also have been encouraged because readings were recommended for each lecture, and therefore some students probably read the e-book in preparation for lecture and later revisited those materials prior to quizzing or exams. Courses with different incentive structures would likely yield different patterns of study, and this may limit the generalizability of our findings.

Further, the relationship between spacing and performance may not generalize to courses in which cramming is a more viable option. Previous research has shown that cramming can sometimes produce good test scores when test delays are short (see Rohrer, 2015). In many college courses (including the one examined here), students can study until the start of an exam -- i.e., a very short test delay. If students are primarily concerned with scores on an immediate test rather than long-term retention, they may find cramming to be an effective strategy. However, our results nonetheless suggest a benefit of spaced study in this context. One possibility is that students may not be able to cram all information in a college course when faced with real-world time constraints and the quantity and complexity of collegiate-level material. Further, cramming may be less effective with application and analysis questions that require deeper understanding, which are common in many college courses. Thus, spaced practice may still benefit exam performance even when long-term retention is ostensibly not required. Future investigations might explore the possibility that cramming is less viable when to-be-learned materials are more complex or interrelated or require deeper understanding (e.g., inference, application).

4.2. Implications

The present research suggests that self-selected spaced study (especially when involving retrieval practice) is consequential for real-world student performance. It also highlights the key role that instructors can play in encouraging beneficial study behaviors. Here, we saw that course structure, deadlines, and incentives influenced students' study behavior. To promote long-term retention of important material, instructors could incentivize spacing of retrieval practice across the duration of a course. For example, frequent low-stakes quizzing is one way for teachers to promote spaced retrieval practice in their classrooms. Alternatively, students could earn points for completing practice quizzes multiple times before an exam. Perhaps, to earn full credit, students might be required to quiz themselves on three separate days and begin at least two weeks before an exam. While students can sometimes benefit from controlling their own study (e.g., Tullis et al., 2018), students also frequently make bad study decisions, such as waiting to study until an exam approaches or not taking advantage of study tools offered to them. By using low-stakes incentives – minimal points per study session which accumulate to be meaningful towards the course grade – instructors can shape study behavior without removing student autonomy.

References

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

Ariel, R., Dunlosky, J., & Bailey, H. (2009). Agenda-based regulation of study-time allocation: When agendas override item-based monitoring. *Journal of Experimental Psychology: General*, 138(3), 432-447. https://doi.org/10.1037/a0015928

Bahrick, H. P. (1979). Maintenance of knowledge: Questions about memory we forgot to ask. *Journal of Experimental Psychology: General*, *108*(3), 296-308. https://doi.org/10.1037/0096-3445.108.3.296

Bartoszewski, B. L., & Gurung, R. A. (2015). Comparing the relationship of learning techniques and exam score. *Scholarship of Teaching and Learning in Psychology, 1*(3), 219. https://doi.org/10.1037/stl0000036

Barzagar Nazari, K., & Ebersbach, M. (2018). Distributed practice: Rarely realized in self-regulated mathematical learning. *Frontiers in Psychology, 9*, 2170-2183. https://doi.org/10.3389/fpsyg.2018.02170

Benjamin, A. S., & Bird, R. D. (2006). Metacognitive control of the spacing of study repetitions. *Journal of Memory and Language*, *55*(1), 126-137. https://doi.org/10.1016/j.jml.2006.02.003

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

Bjork, R. A. (1994). Memory and metamemory considerations in the training of human beings. In J. Metcalfe and A. Shimamura (Eds.), *Metacognition: Knowing about knowing* (pp. 185–205). MIT Press. https://doi.org/10.7551/mitpress/4561.003.0011

Blasiman, R. N., Dunlosky, J., Rawson, K. A. (2017). The what, how much, and when of study strategies: Comparing intended versus actual study behaviour. *Memory, 25*(6), 784-792. https://doi.org/10.1080/09658211.2016.1221974

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. https://doi.org/10.1002/acp.1507

Cepeda, N. J., Pashler, H., Vul, E., Wixted, J. T., & Rohrer, D. (2006). Distributed practice in verbal recall tasks: A review and quantitative synthesis. *Psychological Bulletin, 132*(3), 354-380. https://doi.org/10.1037/0033-2909.132.3.354

Cohen, M. S., Yan, V. X., Halamish, V., & Bjork, R. A. (2013). Do students think that difficult or valuable materials should be restudied sooner rather than later? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 39*(6), 1682-1696. https://doi.org/10.1037/a0032425

Delaney, P. F., Verkoeijen, P. P., & 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. https://doi.org/10.1016/S0079-7421(10)53003-2

Dempster, F. N. (1989). Spacing effects and their implications for theory and practice. *Educational Psychology Review, 1*(4), 309-330. https://doi.org/10.1007/BF01320097

Dunlosky, J., & Rawson, K. A. (2015). Do students use testing and feedback while learning? A focus on key concept definitions and learning to criterion. *Learning and Instruction*, *39*, 32-44. https://doi.org/10.1016/j.learninstruc.2015.05.003

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

Gašević, D., Dawson, S., Rogers, T., & Gasevic, D. (2016). Learning analytics should not promote one size fits all: The effects of instructional conditions in predicting academic success. *The Internet and Higher Education, 28*, 68-84. https://doi.org/10.1016/j.iheduc.2015.10.002

Gilovich, T., Keltner, D., Chen, S., & Nisbett, R. E. (2015). *Social Psychology* (4th ed.). New York: W.W. Norton & Company.

Greene, J. A., & Azevedo, R. (2010). The measurement of learners' self-regulated cognitive and metacognitive processes while using computer-based learning environments. *Educational Psychologist, 45*(4), 203-209. https://doi.org/10.1080/00461520.2010.515935

Hartwig, M. K., & Dunlosky, J. (2012). Study strategies of college students: Are self-testing and scheduling related to achievement? *Psychonomic Bulletin and Review, 19*(1), 126-134. https://doi.org/10.3758/s13423-011-0181-y

Hopkins, R. F., Lyle, K. B., Hieb, J. L., & Ralston, P. A. (2016). Spaced retrieval practice increases college students' short- and long-term retention of mathematics knowledge. *Educational Psychology Review*, *28*(4), 853-873. https://doi.org/10.1007/s10648-015-9349-8

Jansen, R. S., van Leeuwen, A., Janssen, J., Conijn, R., & Kester, L. (2020). Supporting learners' self-regulated learning in Massive Open Online Courses. *Computers & Education, 146*, 103771. https://doi.org/10.1016/j.compedu.2019.103771

Jovanovic, J., Mirriahi, N., Gašević, D., Dawson, S., & Pardo, A. (2019). Predictive power of regularity of pre-class activities in a flipped classroom. *Computers & Education, 134*, 156-168. https://doi.org/10.1016/j.compedu.2019.02.011

Karpicke, J. D., Butler, A. C., & Roediger, H. L., III. (2009). Metacognitive strategies in student learning: Do students practice retrieval when they study on their own? *Memory, 17*(4), 471-479. https://doi.org/10.1080/09658210802647009

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

Kornell, N., & Bjork, R. A. (2007). The promise and perils of self-regulated study. *Psychonomic Bulletin & Review, 14*(2), 219-224. https://doi.org/10.3758/BF03194055

Kuhbandner, C., & Emmerdinger, K. J. (2019). Do students really prefer repeated rereading over testing when studying textbooks? A reexamination. *Memory, 27*(7), 952-961. https://doi.org/10.1080/09658211.2019.1610177

Michael, J. (1991). A behavioral perspective on college teaching. *The Behavior Analyst*, *14*, 229–239. https://doi.org/10.1007/BF03392578

Morehead, K., Rhodes, M. G., & DeLozier, S. (2016). Instructor and student knowledge of study strategies. *Memory, 24*(2), 257-271. https://doi.org/10.1080/09658211.2014.1001992

Moulton, C. A. E., Dubrowski, A., MacRae, H., Graham, B., Grober, E., & Reznick, R. (2006). Teaching surgical skills: What kind of practice makes perfect? A randomized, controlled trial. *Annals of surgery*, *244*(3), 400-409. https://doi.org/10.1097/01.sla.0000234808.85789.6a

Nandagopal, K., & Ericsson, K. A. (2012). An expert performance approach to the study of individual differences in self-regulated learning activities in upper-level college students. *Learning and Individual Differences, 22(*5), 597-609. https://doi.org/10.1016/j.lindif.2011.11.018

Nisbett, R. E., & Wilson, T. D. (1977). Telling more than we can know: Verbal reports on mental processes. *Psychological Review, 84*(3), 231-259. https://doi.org/10.1037/0033-295X.84.3.231

Plant, E. A., Ericsson, K. A., Hill, L., & Asberg, K. (2005). Why study time does not predict grade point average across college students: Implications of deliberate practice for academic performance. *Contemporary Educational Psychology, 30*(1), 96-116. https://doi.org/10.1016/j.cedpsych.2004.06.001

Pyc, M. A., & Dunlosky, J. (2010). Toward an understanding of students' allocation of study time: Why do they decide to mass or space their practice? *Memory & Cognition, 38*(4), 431-440. https://doi.org/10.3758/MC.38.4.431

Rawson, K. A., Vaughn, K.E., & Carpenter, S. K. (2015). Does the benefit of testing depend on lag, and if so, why? Evaluating the elaborative retrieval hypothesis. *Memory & Cognition, 43*(4), 619-633. https://doi.org/10.3758/s13421-014-0477-z

Roediger III, H. L., & Karpicke, J. D. (2006). The power of testing memory: Basic research and implications for educational practice. *Perspectives on Psychological Science*, *1*(3), 181-210. https://doi.org/10.1111/j.1745-6916.2006.00012.x

Rohrer, D. (2015). Student instruction should be distributed over long time periods. *Educational Psychology Review, 27*(4), 635-643. https://doi.org/10.1007/s10648-015-9332-4

Romero, C., & Ventura, S. (2020). Educational data mining and learning analytics: An updated survey. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, 10*(3), 1-21. https://doi.org/10.1002/widm.1355

Seabrook, R., Brown, G. D. A., & Solity, J. E. (2005). Distributed and massed practice: From laboratory to classroom. *Applied Cognitive Psychology*, *19*(1), 107-122. https://doi.org/10.1002/acp.1066

Son, L. K. (2004). Spacing one's study: evidence for a metacognitive control strategy. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 30*(3), 601-604. https://doi.org/10.1037/0278-7393.30.3.601

Sobel, H. S., Cepeda, N. J., & Kapler, I. V. (2011). Spacing effects in real-world classroom vocabulary learning. *Applied Cognitive Psychology*, *25*(5), 763-767. https://doi.org/10.1002/acp.1747

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. https://doi.org/10.1007/s11251-012-9231-8

Taraban, R., Maki, W. S., & Rynearson, K. (1999). Measuring study time distributions: Implications for designing computer-based courses. *Behavioral Research Methods, Instruments, and Computers, 31*(2), 263-269. https://doi.org/10.3758/BF03207718

Toppino, T. C., & Cohen, M. S. (2010). Metacognitive control and spaced practice: Clarifying what people do and why. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *36*(6), 1480-1491. https://doi.org/10.1037/a0020949

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*(4), 540-552. https://doi.org/10.1037/xlm0000473.

van Leeuwen, A., Bos, N., van Ravenswaaij, H., & van Oostenrijk, J. (2019). The role of temporal patterns in students' behavior for predicting course performance: A comparison of two blended learning courses. *British Journal of Educational Technology*, *50*(2), 921-933. https://doi.org/10.1111/bjet.12616

Appendix A

Table A1

Study Time Questions

Questions	Response Format	
Approximately how many hours & minutes did you spend (outside of lecture) studying for this class since the last exam (adding up all the time you spent reading the textbook, completing quiz tool assignments, studying your notes, and studying the material in any other way)?	Approximately ______ hours & ______ minutes	
When you think about the time you spent studying for this class (outside of lecture) since the last exam, how much of that time was spent doing the following activities? (Altogether, this should total 100% of your study time.)	<u>Activity</u>	<u>% of your study time</u>
	Reading/studying the e-book	__________%
	Using the quiz tool	__________%
	Studying your lecture notes	__________%
	Other (if applicable): _________	__________%

Note. These two questions were asked immediately following Exams 4 and 5. Responses indicated that e-book and quiz tool usage constituted more than half of students' total study time.

Appendix B

Table B1

Predicting Student Performance on Each Exam from Indicators of Spacing (Frequency, First Use) While Controlling for Time

				Regression for Frequency		Regression for First Use	
Study tool	Exam	Required Chapter	Sample size (N)	Predictor β (frequency)	Control β (tool time)	Predictor β (first use)	Control β (tool time)
Quiz tool							
	1	1	173	.29**	-.25**	.25**	-.16*
	2	3	169	.20*	-.13	.22**	-.07
	3	5	174	.16†	-.08	.22**	-.02
	4	7	157	.15†	-.20*	.13	-.16*
	5	6	166	.19*	-.15	.16*	-.07
E-book							
	1	1	173	.04	-.03	.08	-.03
	2	3	157	.11	-.19	.21*	-.18*
	3	5	142	-.13	.17	.13	.04
	4	7	122	-.10	.14	.11	.02
	5	6	128	.17	.01	.07	.14

**$p < .01$ *$p < .05$ †$p < .10$

Table 1

Frequency of Online Retrieval Practice and Reading

		Quiz tool (retrieval practice)										E-book (reading)													
		# Non-users	# Users	Frequency (number of days used)							Mean freq.	# Non-users	# Users	Frequency (number of days used)											Mean freq.
Exam	Chapter			1	2	3	4	5	6	7				1	2	3	4	5	6	7	8	9	>9		
	Required																								
1	1	8	173	125	34	9	3	1	1		1.3	8	173	40	68	28	19	10	5	2	1			2.4
2	3	6	169	123	36	7	2	1			1.3	18	157	33	55	34	14	11	4	1	2	1	2	2.5
3	5	6	174	122	44	8					1.3	38	142	56	40	25	15	4		2				1.7
4	7	5	157	115	34	5	2			1	1.3	40	122	42	42	19	9	4	4		2			1.7
5	6	10	166	134	27	4		1			1.2	48	128	46	46	23	5	6	1	1				1.5
	Mean %	***4.0***	***96.0***	***70.9***	***20.0***	***3.8***	***0.8***	***0.3***	***0.1***	***0.1***	***-***	***17.6***	***82.4***	***24.8***	***28.7***	***14.7***	***7.0***	***4.0***	***1.6***	***0.7***	***0.6***	***0.1***	***0.2***	***-***
	Optional																								
1	2	114	67	43	18	4	2				0.5	70	111	65	33	9	2	1	1					1.0
2	4	93	82	64	12	6					0.6	76	99	37	37	11	7	2	4	1				1.2
3	11	119	61	52	7	2					0.4	83	97	59	26	9	2	1						0.8
4	8	118	44	36	8						0.3	62	100	68	22	9	1							0.9
4	9	122	40	36	4						0.3	92	70	52	15	3								0.6
5	10	139	37	30	5	1	1				0.3	86	90	49	29	8	3			1				0.9
	Mean %	***68.2***	***31.8***	***25.1***	***5.2***	***1.2***	***0.3***	***0***	***0***	***0***	***-***	***45.4***	***54.6***	***32.0***	***15.5***	***4.7***	***1.4***	***0.4***	***0.5***	***0.2***	***0***	***0***	***0***	***-***

Note. Higher frequencies indicate that students spaced their study into more sessions. *Required* indicates that using the quiz tool was required for that chapter, and *optional* indicates that using the quiz tool was optional for that chapter (though students knew that exams would cover all eleven chapters). Our total sample size was 185 students, but not every student took every exam. Here, we show data for the students who took the exams (181 students for Exam 1; 175 for Exam 2; 180 for Exam 3; 162 for Exam 4; and 176 for Exam 5). The few untested students per exam were excluded because their frequencies would not be representative of their normal behavior.

Table 2

Predicting Student Performance from Indicators of Spacing (Frequency, First Use) While Controlling for Time

				Correlations		Regression Analyses			
						DV: Course perf.		DV: College GPA	
Variables/Predictors	*Mean*	*SD*	*N*	Course perf.	College GPA	Predictor β	Control β (time)	Predictor β	Control β (time)
Performance									
Course performance	75.7	11.9	185	--	.53**	--	--	--	--
College GPA	3.1	0.5	167	.53**	--	--	--	--	--
Quiz tool			185						
Frequency	1.3	0.5		.19**	.21**	.37**	-.33**	.36**	-.28**
First use	10.2	2.4		.26**	.35**	.31**	-.21**	.38**	-.16*
Tool time (min)	34.4	14.0		-.13	-.08	--	--	--	--
E-book			182						
Frequency	2.0	1.2		.01	.05	.03	-.03	.06	-.01
First use	8.6	4.0		.14†	.15†	.20*	-.11	.18*	-.06
Tool time (min)	102.2	74.2		-.00	.04	--	--	--	--

Note. Descriptive statistics and bivariate correlations are shown on the left, and regression analyses are shown on the right. Separate regression models were run for each dependent variable (DV) and predictor. Three students never used the e-book during the semester and were therefore excluded from e-book analyses (N=182). Analyses involving GPA had a smaller sample size (*N*=167) because 18 students did not provide their GPA.
**$p < .01$ *$p < .05$ † $p = .054$

Figure 1

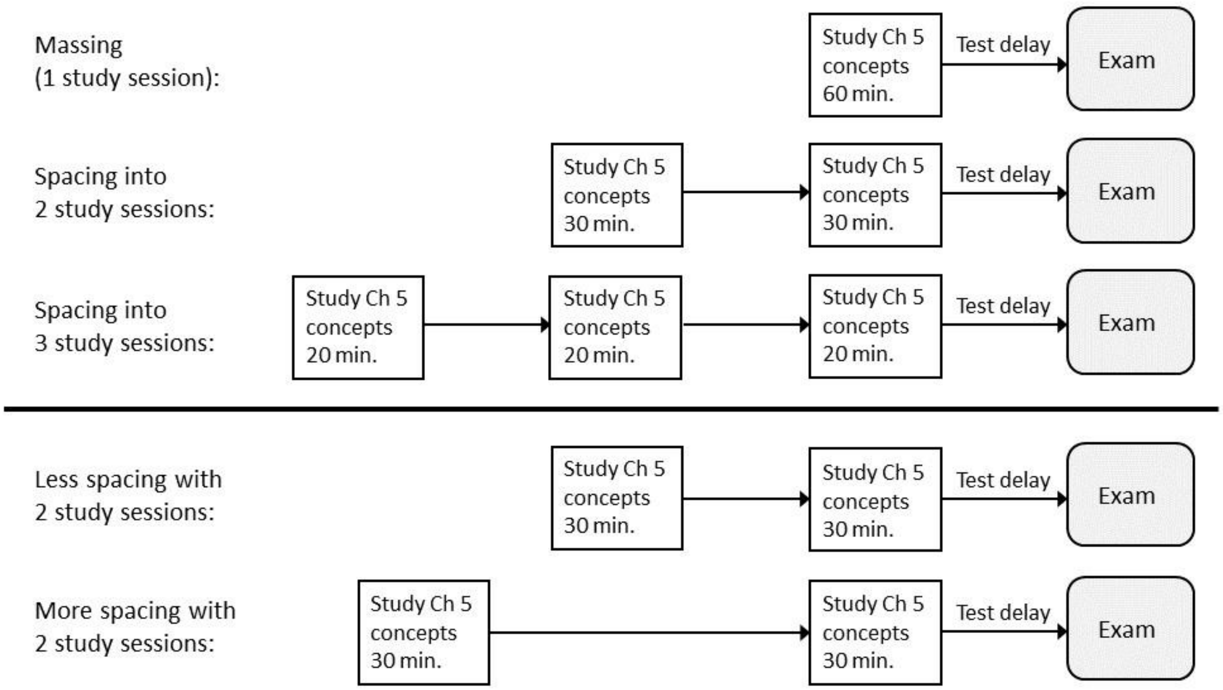

Figure 1. Examples of how college students might space their study for a course exam. The top panel shows how spacing can be increased by spreading study across more sessions, while total study time remains the same. The bottom panel shows how spacing can be increased by spreading study sessions across a longer span of time, without changing the number of sessions or time spent studying.

Figure 2

	Jan 11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27
															D		Exam #1
Req. **QUIZ TOOL**								15	6	12	16	12	9	40	87	24	25
Opt.								1	1	2	1	4	1	4	17	33	35
Req. **E-BOOK**								75	22	24	16	25	16	46	92	84	57
Opt.								5	3	5	2	8	6	13	26	59	51

	28	29	30	31	Feb 1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22
												D											*			#2
Req. **QUIZ TOOL**	1	-	-	-	1	1	20	15	11	8	23	111	5	-	-	-	-	-	-	2	2	2	2	-	13	28
Opt.	-	-	-	-	-	1	-	-	-	-	-	-	1	1	-	-	-	-	6	9	3	7	12	2	31	37
Req. **E-BOOK**	7	1	1	2	2	5	29	14	14	9	22	95	6	1	2	2	2	2	11	22	28	23	21	13	54	73
Opt.	2	-	-	-	-	1	3	3	2	2	2	1	1	-	-	1	1	1	10	25	33	19	17	10	34	53

	23	24	25	26	27	28	29	Mar 1	2	3	4	5	6	7	8	9
							(D)		D							#3
Req. **QUIZ TOOL**	10	10	9	8	8	24	78	10	45			1	-	2	9	25
Opt.	1	-	-	-	-	1	-	-	-			3	1	7	18	41
Req. **E-BOOK**	20	15	12	5	6	19	54	11	28			18	4	23	57	48
Opt.	1	2	3	1	-	2	3	3	4			15	4	20	49	53

	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	Apr 1	2	3	4
					Spring Break														D							#4
Req. **QUIZ TOOL**		1	-	-	2	1	2	1	-	2	5	16	14	15	10	10	8	22	104	3	2	2	-	1	5	15
Opt.		-	-	-	-	-	-	-	-	-	-	-	-	-	1	-	-	-	1	1	4	2	1	4	18	28
Opt.		-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	1	-	2	2	14	28
Req. **E-BOOK**		6	-	1	1	-	-	1	1	1	3	19	13	14	12	9	7	17	52	5	7	9	8	7	51	66
Opt.		4	-	-	-	-	-	-	-	-	-	1	4	1	3	1	-	-	3	4	5	6	6	6	39	75
Opt.		2	-	-	-	-	-	-	-	-	-	-	1	-	-	-	-	-	1	1	4	7	2	3	30	48

	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22
									D									#5
Req. **QUIZ TOOL**	9	11	8	5	1	8	34	25	87	1	-	-	1	1	2	1	5	17
Opt.	-	-	-	-	-	-	-	1	1	1	-	1	1	2	3	4	11	23
Req. **E-BOOK**	12	10	4	2	2	9	25	17	43	1	1	1	5	15	14	24	41	54
Opt.	7	-	-	1	-	1	4	3	2	1	1	1	1	9	12	10	39	62

Figure 2. Course schedule. The semester began on Jan 11; recording of tool usage began on Jan 18. (Dates shaded in gray indicate usage was successfully recorded. Downloads were unsuccessful on 3 days: March 3, 4, and 10. Any usage on those days was included in the next successful download, e.g., March 11 is the cumulative usage across March 10 and 11.) Exams occurred on Jan 27, Feb 22, Mar 9, Apr 4, and Apr 22. (Exam 2 was originally scheduled for Feb 19 but was switched by the instructor to Feb 22.) Deadlines for required quiz tool assignments are denoted with a D. (The deadline for the quiz tool assignment for Exam 3 was extended by the instructor from Feb 29 to Mar 2.) Beneath each date, we show how many students (out of 185) used the quiz tool or e-book on that date, separately for each required or optional chapter.

Figure 3

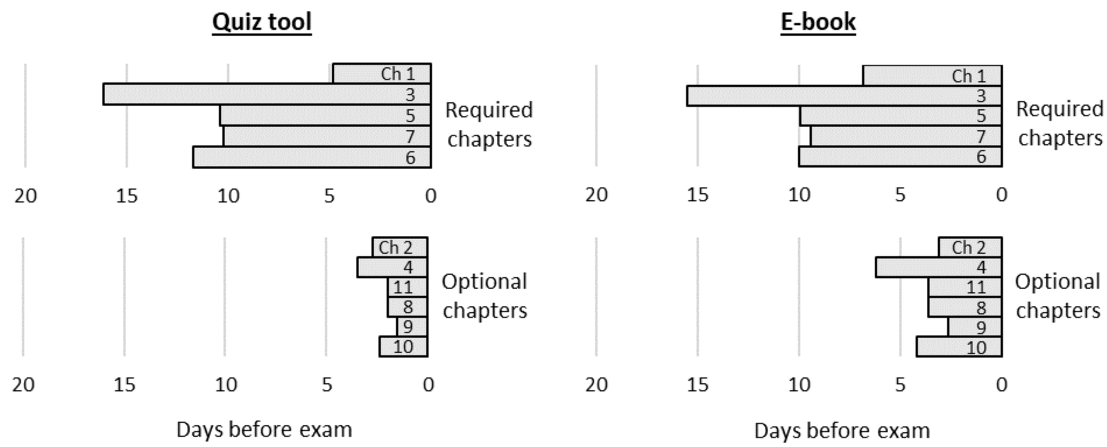

Figure 3. How long before an exam (in days) a tool was first used. Studying that begins sooner is spaced across a longer span of time.