Running Head: SIMULTANEOUS STUDY DECISIONS

Simultaneous decisions at study: Time allocation, ordering, and spacing

Lisa K. Son

Barnard College

Nate Kornell

Williams College

In press at Metacognition and Learning. Please do not quote without permission.

Abstract

Learners of all ages face complex decisions about how to study effectively. Here we investigated three such decisions made in concert - time allocation, ordering, and spacing. First, college students were presented with, and made judgments of learning about, 16 word-synonym pairs. Then, when presented with all 16 pairs, they created their own study schedule by choosing when and how long to study each item. The results indicated that a) the most study time was allocated to difficult items, b) relatively easy items tended to be studied first, and c) participants spaced their study at a rate significantly greater than chance. The spacing data, which are of particular interest, differ from previous findings that have suggested that people, including adults, believe massing is more effective than spacing.

Key Words: Spacing, Metacognition, Studying, Learning, Flashcards

Simultaneous decisions at study: Time allocation, ordering, and spacing

Learners are faced with a variety of decisions to make during study. Consider a student studying a set of words for a vocabulary test, or preparing for upcoming exams in math, science, and history. One decision is to *allocate* study time to each subject. For example, one might allocate the most time to the subject that is judged to be close to, but not yet fully, learned (e.g. Kornell & Metcalfe, 2006). In addition, the student would need to decide the *order* in which to study the materials. For instance, the learner might want to begin by studying the easy rather than the more difficult materials (e.g. Thiede & Dunlosky, 1999). A third decision, if one studies material repeatedly, is to choose the *spacing*, or the amount of time that should elapse, before re-study (e.g. Son, 2004). The current experiment investigates the manner in which students unravel, simultaneously, these three study decisions: time allocation, ordering, and spacing.

Research on metacognition — beliefs and judgments about memory — has shown that people can use their metacognitive knowledge to control their study decisions in a variety of tasks (Nelson & Narens, 1990). Below, we summarize relevant findings on the use of metacognition to control time allocation, ordering, and spacing.

Suppose that a student studying for exams were faced with only time allocation decision. That student could allocate time based on how well learned a topic was. According to the *discrepancy-reduction model* (*DR*) (Dunlosky & Hertzog, 1998), people spend the most time on items that are furthest from the desired state. For example, if math were the furthest from being learned, then the most study time would be allocated to math. This allocation strategy has, until recently, been consistent with virtually all research on study-time allocation (see Son & Kornell, 2008, for a review). Recent research has shown, however, that under time pressure and competition for resources, *DR's* predictions can be inaccurate. Another model, the *region of proximal learning model (RPL)*, predicts that in such situations a more effective strategy—and the strategy that people use—is to allocate time to the relatively easy items, as long as they are not already learned (e.g. Metcalfe & Kornell, 2003, 2005).

Now suppose that the student faces the problem of ordering his or her study. Given the abundant data showing that people studied the difficult items the longest, it seems reasonable to assume that those items would also be studied first (Dunlosky & Hertzog, l998). A few studies have suggested that, on the contrary, people choose to study the *easier* items first, and then shift to the difficult items afterwards (e.g. Thiede & Dunlosky, 1999). Metcalfe & Kornell (2005), for example, simultaneously presented nine items, in a 3 X 3 array, for participants to study as they wished. Of the nine items, the left-most items were relatively easy, the middle were medium, and the right-most were relatively difficult. The data showed that people studied the easy, medium, and difficult items in that order. Although people might have used position to make their choices, a previous study (Metcalfe and Kornell, 2003) found that when people were asked to select easy, medium, and difficult items for study, they tended to study the easy items first when very time was short; as time availability increased, study was increasingly shifted to the medium and difficult items.

The third, and perhaps the most complicated, decision students face—for topics that are studied more than once—is how far to space out the study of one particular topic. For instance, suppose that the learner had 6 hours in which to study math, history, and science. The learner could mass study: Study math for 2 hours, then history for the next 2 hours, and finally science for the remaining 2 hours. Or, he could use a spacing schedule: Spend an hour on each topic for the first 3 hours, and then repeat the cycle during the last 3 hours. Although in both scenarios the amount of time allocated to each topic is the same, research has shown that spacing study is more effective than massing study—a phenomenon known as the *spacing effect* (Glenberg, 1979; Dempster, 1989; Cepeda, Pashler, Vul, Wixted, & Rohrer, 2006).

Despite the positive effects of spaced study, when asked to make judgments of how much they learn under different study schedules, people often give higher ratings following massed than spaced practice (Kornell & Bjork, 2008; Simon & Bjork, 2001; Zechmeister & Shaughnessy, 1980). Massed practice may seem effective because it requires less effortful processing than spaced practice, and because massing makes an item extremely fluent—and thus seem learned—on the second presentation.

The fact that the spacing effect is one of the most robust effects in cognitive psychology (Dempster, 1989), combined with the fact that students seem to view spacing as harmful, not helpful, point to a discouraging prediction: When choosing how to schedule their spacing, students will choose to mass, and in doing so seriously undermine the effectiveness of their study.

Despite the importance of using spaced study schedules, and the seeming likelihood that people will fail to do so, very little is known about how people schedule their study. One study found that scheduling choices are systematically related to metacognitive judgments (Son, 2004). That is, when people were asked whether they would like to mass or space their study, they choose neither to mass exclusively nor to space exclusively, but rather to mass items they judged as difficult and space items they judged as easier. In a different paradigm, however, in which participants were not given completely free choices to mass or space, and instead were limited to choosing to mass half of the items and space the other half, the opposite pattern was obtained (Benjamin & Bird, 2006). Participants massed the easy items and spaced the harder items. Given that the scheduling strategies people use vary from situation to situation, we attempted to develop an experimental paradigm that mimics real-world study, to achieve an understanding of how and why people space and mass their study.

In the real world, people rarely have the luxury of making isolated decisions about time allocation, ordering, and spacing—instead, they have to make decisions about all three simultaneously. To our knowledge, no research has investigated how the three strategies are processed in combination. We combined these three decisions, in the current experiment, for two reasons: First, in order to better understand how the three types of decision relate both to each other and to metacognitive judgments, and second, to examine study decisions in an ecologically valid situation.

Of course there are limits to the ecological validity of almost any empirical study. For instance, students typically have the opportunity to study across days, weeks, or months (although they do not necessarily take advantage of that opportunity, but, instead, tend to cram just before they are tested), the current study took place in a single day. Nevertheless, a controlled experiment, even within one study session, can unveil the basic strategies that people use when learning. Another departure from reality might be that people do not make *overt* judgments of learning during study. In this experiment, though, we explicitly asked students to make conscious assessments about each of the to-be-learned items. This procedure allows for us to directly test where breakdowns occur during *subsequent* decision making. Although there are certainly other factors that affect study decisions, including goals, social situations, working memory capacity, the objectives of the current experiment are twofold: a) To examine, in a realistic situation, the study strategies people use when facing the triple problem of allocation, ordering, and spacing, and b) to understand the role of metacognitive knowledge in such strategy choices.

Method

We used a method in which participants controlled allocation, order, and spacing. Participants were presented with 16 word-synonym pairs for study, and then, for each pair, made a Judgment of Learning (JOL)—that is a judgment about the likelihood that they would be able to remember the pair on a later test. Then all of the items were presented simultaneously, and people were asked to create their own study list—which was 8, 16, or 24 items long depending on the condition—by choosing and organizing the items they wanted to study. Critically, they could choose to repeat any of the items, and if they so desired, could choose *not* to study any given item. Thus participants could control both the distribution of items (i.e., the order and spacing) and the number of times to study each item (i.e., time allocation).

Participants.

Seventy-three college students participated for course credit. They were recruited from their Introductory Psychology courses, where participation in a variety of experimental sessions was required. No other incentives were offered except for the fact that they were told that they would be taking a memory test, and to do the best they could. There were 25, 24, and 24 participants in the 8, 16, and 24 restudy trial conditions, respectively.

Materials.

The stimuli were 16 word-synonym pairs (e.g., *ignominious-shameful*), randomly selected by the computer from a pool of 100 pairs, taken from the Graduate Record Examination. We allowed for a range in difficulty and assumed that the level of difficulty for each pair would vary from individual to individual. The critical measure of difficulty, then, was a participant's individual judgment of learning for each item.

Design.

Participants were allowed 8, 16, or 24 restudy trials in a between-participants design. Restudy trials will be referred to as “slots” because when creating their study schedule, participants were provided with, for example, 8 empty slots in which to place items for restudy.

Procedure.

Participants were first presented with a sequence of 16 word-synonym pairs to study for a later test. Each pair was presented for 2 seconds—one second of the word, and one second of the synonym. After a pair was presented, participants made a JOL, that is, they made a rating, by typing a number from 0-10, indicating how confident they were that they would be able to recall the synonym, when given only the word, on a later memory test.

After making their JOLs, all 16 cues (without targets) were shown simultaneously on the left side of the computer screen. On the right of the screen was a list of slots to be filled for re-study. The number of slots on the list varied by condition—participants either saw 8, 16, or 24 slots (in conditions hereafter named *8slot, 16slot, 24slot,* respectively*)*. In *8slot*, at most half of the 16 items could be restudied. The participant's task was to click on a cue that they wanted to re-study and drag it into one of the slots.

Participants were told that after a word had been placed in each slot, they would study the pairs from top to bottom, and thus, for instance, if one wanted to study a particular item first and third during the re-study phase, then the item should be placed into slots 1 and 3. They were also told that they could study pairs as many (or few) times as they wanted. For instance, a participant could study a particular item zero times, and another 3 times, and those three could be distributed freely. Thus, participants fully controlled the amount of time and position of every item. The only constraint was that all of the slots had to be filled.

Once the participant had created a complete study list, there was a re-study period. During this period, the participant saw the cue-target pairs sequentially, in the exact order that he or she had chosen, for 3 seconds each. After a 3-minute distractor task, during which participants completed multiplication problems, a final cued-recall test was conducted. Participants were presented with each of the 16 cues, one at a time in a random order, and their task was to type in the synonym.

Results

Participants made JOLs on a scale of 0-10 during the initial study phase. The mean JOLs did not differ significantly between the *8slot* ($M = 5.77$, $SD = 1.64$), *16slot* ($M = 5.23$, $SD = 1.46$), and *24slot* ($M = 5.47$, $SD = 1.34$) conditions, $F < 1$. The participants reported a wide range of JOLs. We computed a JOL range for each participant separately. The mean low JOL was 1.99, the mean high was 9.08 and the mean range was 7.10.

The JOLs made during the initial study phase accurately predicted final test performance. A Goodman-Kruskal Gamma correlation between JOL and final test accuracy was computed for each participant; the mean correlation was significantly above zero for *8slot* ($M = .50$, $t(23) = 5.98$, $p < .0001$, $d = 1.22$), *16slot* ($M = .49$, $t(19) = 6.30$, $p < .0001$, $d = 1.41$), and *24slot* ($M = .46$, $t(19) = 5.94$, $p < .0001$, $d = 1.33$). The groups did not differ significantly, $F < 1$.[1] Aside from their accuracy, it is difficult to draw conclusions based on JOLs, because the study session that took place between the JOL phase and the final test may have had important consequences that the participants did not necessarily anticipate while making JOLs. The critical issue here, however, is how people's JOLs guided their study strategies regarding time allocation, ordering, and spacing.

Time allocation

Correlations were computed between an item's JOL and the number of times it was scheduled for study. The correlations were significantly negative for *8slot* ($M = -.37$, $t(24) = 3.96$, $p < .0001$, $d = .79$), *16slot* ($M = -.53$, $t(23) = 8.57$, $p < .0001$, d = 1.75), and *24slot* ($M = -.55$, $t(23) = 9.13$, $p < .001$, d = 1.86). These means did not differ significantly, $F(2,70) = 1.86$, p = .16. Thus in all three conditions, participants focused their time on the relatively judged-difficult items. The data were also analyzed by categorizing each item as either High or Low JOL, based on a median split of JOLs for each participant. As Figure 1 shows, participants studied relatively difficult (i.e., low

---

[1] Gamma correlations could not be computed for one, four, and four participants in *8slot*, *16slot*, and *24slot*, respectively, because their accuracy was perfect.

JOL) items more frequently than easy (i.e., high JOL) items, $F(1,70) = 76.86$, $p < .0001$, $\eta_p^2 = .052$. By definition there was an effect of group; there was also a significant interaction between group and JOL level, $F(2,70) = 3.27$, $p = < .05$, $\eta_p^2 = .09$, probably driven by the larger range of possible study times in *16slot* and *24slot* than in *8slot*.

Ordering

Although participants chose to spend the most time studying difficult items, those items were not necessarily studied first. There were marginally significant correlations between JOL and the slot number, showing that participants placed easier items in relatively early slots, in *8slot* ($M = -.15$, $t(22) = 1.89$, $p = .07$, $d = .39$) and *16slot* ($M = -.16$, $t(21) = 1.84$, $p = .08$, $d = .39$). The correlation was not significant in the *24slot* ($M = -.01$). There was not a significant difference between the means in the three groups ($F < 1$). Though small, these effects are consistent with *RPL's* prediction that, when study opportunities are limited, people will study easy items first, and then turn to harder items. One explanation for why the correlations are small is that, instead of studying the easiest items first, participants often chose not to study them at all, so that all of the items that were studied were relatively difficult.

Spacing

Spacing requires that an item be studied at least twice. Here, we only analyzed items that were studied exactly twice. Items studied three or four times were excluded[2]. There were two reasons for this exclusion rule. First, it seemed impossible to accurately quantify, for example, which has greater spacing: an item studied three times at uniform intervals, or an item studied with larger total spacing between the first and third

---

[2] No participant ever studied an item more than four times.

repetition, but with a very short second interval. Second, no items were studied more than twice in *8slot*, and in *16slot* and *24slot*, only 4% and 14% of items were studied more than twice, respectively. Items studied twice accounted for 4%, 25% and 36% of the items in *8slot*, *16slot* and *24slot*, respectively.

To quantify the degree to which participants spaced their practice, we computed the total spacing (i.e., number of intervening items) from the first to second presentation of an item. For example, if an item was placed in slot 2 and 4, its spacing was one, because there was one intervening item. We then computed an average spacing across items.

To test the hypothesis that participants spaced their study, we compared each participant's average spacing to the amount of spacing that would be expected if items were arranged randomly. An expected random spacing score was computed, separately for each participant, by simulating random distributions of items given the number of items the participant studied twice and the number of slots they were allotted.

Results showed that participants spaced their practice significantly more than would be expected by chance. As Figure 2 shows, average spacing exceeded the amount of spacing expected from a random distribution of items, $F(1,52) = 16.56$, $p < .001$, $\eta_p^2 = .24$. There was also a significant effect of group; not surprisingly, spacing was greater when more slots were available, $F(2,52) = 69.35$, $p < .0001$, $\eta_p^2 = .73$. The two variables did not interact, $F(2,52) = 2.28$, $p = .11$. We also compared actual spacing to the maximum possible spacing (i.e., the amount of spacing that would result if all items studied more than twice were studied as close to the beginning and end of the list as possible). Actual spacing was significantly smaller than maximum spacing, $F(1,52) =$

73.36, $p < .0001$, $\eta_p^2 = .59$. Again, spacing was greater when there were more slots available, $F(2,52) = 95.87$, $p < .0001$, $\eta_p^2 = .79$. There was also a significant interaction between the two variables, $F(2,52) = 3.97$, $p < .05$, $\eta_p^2 = .13$, presumably because there was a greater difference between actual and maximum spacing in *16slot* and *24slot* than in *8slot*, which may, again, have occurred due to the restricted range in *8slot*.

Did spacing vary as a function of JOLs? The *8slot* group was excluded from the analysis because only three participants had the necessary number of observations—at least two items, with different JOLs, had to be studied twice—to be analyzed. Nineteen and 21 participants met this requirement in *16slot* and *24slot*, respectively. There was no significant difference in spacing between items given low and high JOLs, $F < 1$, nor did this variable interact with the number of slots allowed for study $F < 1$. Of course, spacing was greater in *24slot* than *16slot*, $F(1,38) = 20.85$, $p < .0001$, $\eta_p^2 = .53$. Thus, in contrast to previous findings (Benjamin & Bird, 2006; Son, 2004), there was no evidence that participants decided how much to space an item based on its difficulty.[3]

Final test performance

The three groups did not differ in final test performance ($F < 1$), despite the difference in the number of restudy opportunities. There was also no significant

---

[3] Items studied three times or four times were counted as intervening items when computing the average spacing of items studied twice. This reflects the spacing choices participants made, but a drawback is that, if participants spaced the triples and quadruples by studying them at the start and end of a list, it would fill those slots and compress the slots remaining for items studied twice. Therefore the data were also analyzed throwing out triples and quadruples completely, so that, for example, if one item was studied three times in *16slot*, it was treated as though that item did not exist and there were only 13 slots. The findings were essentially the same; all significant findings remained significant, and no non-significant findings became significant. There was one exception: The interaction between actual versus random spacing and condition—which was significant in the primary analysis—was only marginally significant, $F(2,52) = 3.15$, $p = .05$.]

correlation between the number of times an item was studied and final test accuracy. In fact, the average correlation was negative (M = -.09, $t(63) = 1.15$, $p = .26$), probably because the most difficult items were selected for study most often.

Correlations were computed for each participant between spacing and final test accuracy. Again, *8slot* was excluded (only one participant had enough observations to be analyzed). In *16slot*, the correlation was significantly greater than zero (M = .68, $t(9) = 3.37$, $p < .01$, $d = 1.07$), which indicated that greater degrees of spacing were correlated with greater memory accuracy. The correlation was not significant in *24slot*, nor was it positive (M = -.17, $t(13) = 1.00$, $p = .34$).

General Discussion

The use of metacognition has been shown to be crucial in boosting learning by allowing learners to make profitable study decisions (e.g. Kornell & Metcalfe, 2006). Here, we focused on three decisions that are often important outside the classroom: time allocation, ordering, and spacing. The results indicated that participants (1) primarily studied items judged to be relatively difficult, (2) tended to study easier items earlier than the more difficult ones, and (3) spaced their study significantly more than would be expected by chance—although the amount an item was spaced was not correlated with its difficulty rating.

The allocation results replicate previous findings—for example, participants in a study by Nelson, Dunlosky, Graf, and Narens (1994) chose to study the most difficult items. Both those and the present data appear to contradict the predictions of *RPL*, namely that under limited time, learners will focus on easier items. In the Nelson et al., study, however, it appears that the participants' main decision was not to seek out difficult items for study per se, but, rather, to avoid studying items that they believed they had already learned (c.f., Kornell & Metcalfe, 2006). The same may have been true here, because many of the relatively easy GRE synonyms may have been familiar enough that participants could remember them after a single study opportunity. If so, the results would be consistent with the predictions of *RPL* as well as *DR*, both of which predict that learners will focus their study on items that they do not yet know. The analyses of study order suggest that, after choosing a set of relatively difficult items to study, participants tended to schedule easier items earlier. This finding provides further support for *RPL*, which predicts that participants will tend to focus on the easiest (unknown) items (Kornell & Metcalfe, 2006).

Spacing

On the surface, the fact that our participants decided to space their study suggests that they may have believed in the benefits of spacing. There is a common misconception, however, that in random sequences, streaks of the same item are rarer than they actually are (Gilovich, Vallone, & Tversky, 1985). Thus, it is also possible that our participants thought that they were distributing items randomly, when they were actually spacing at a rate greater than would be expected by chance.

The fact that participants chose to space their study, whatever their reasons, stands in contrast to the finding that when making JOLs, people seem to consistently believe that massing is more effective than spacing (Kornell & Bjork, 2008; Simon & Bjork, 2001; Zechmeister & Shaughnessy, 1980). These findings may represent a mismatch between metacognitive judgments and study decisions (see Kornell & Son, in press, for a similar disconnect between judgments and choices regarding the benefits of tests as learning events). The present results suggest that people may hold a general belief that massing is not an effective study strategy, and yet, nonetheless, they may also judge massing as more effective than spacing when they are actually engaged in study, because of the strong feeling of fluency that massing creates.

Recent research has produced contradictory evidence about the role of item difficulty in decisions about spacing. Easy items tended to be spaced in one study (Son, 2004), whereas difficult items tended to be spaced in another (Benjamin & Bird, 2006). Although the present experiment was designed, in part, to help resolve this contradiction, the results did not favor either finding. Instead, there was no significant relationship between item difficulty and the amount an item was spaced. In a complex decision environment such as in the present experiment, the relationship between difficulty and spacing may be overshadowed by more fundamental concerns, such as whether to space at all.

Learning outcomes

Although the findings seem to suggest that the people are fairly good at making study decisions—they focused on difficult items, and chose to space out their study trials—assessing the impact of these decisions was difficult because each variable was under the participant's control, and thus only correlational analyses could be conducted. Such analyses are plagued by item-selection effects, such as the fact that participants tend to spend more time studying the items that they found more difficulty initially.

The one variable that was not under the participants' control, number of slots, did not have a significant effect on final test performance: Participants who were provided 24 slots did not significantly outperform participants who only had 8. This surprising finding might be due to the fact that some of the items may have too easily learnable, thus making the allocation, and scheduling, of subsequent study inconsequential. A less intuitive, but perhaps more appropriate, explanation, is that there may have been a *labor-in-vain* effect (Nelson & Leonesio, 1988), where learners allocate too many slots to the too-difficult items, and gain little by doing so (in particular if the 3 additional seconds were insufficient for achieving successful learning). In other words, it may be that the participants were not choosing to study those items in their own *region of proximal learning*. As Kornell & Metcalfe (2006) demonstrated, among items that are unknown, it is most beneficial to focus on items that are *closest* to being learned, rather than items that are more difficult, and perhaps too difficult to learn at all. Thus, a challenge for students is to be savvy enough to know both what they don't know, and in addition, *be able to accurately assess how much it would take to learn the items*. Unfortunately, an issue that has received relatively little attention is how and if people know about an item's *learnability*, including the amount of time it would take to learn. A good guess would be that when people labor in vain, their misjudgments are not due not to a breakdown in knowing what is unknown, but rather, what is and what is not learnable.

Educational implications

With any metacognitive investigation, one would hope that the findings are beneficial to the understanding and/or improvement of learning in the real world. How generalizeable are these data? In laboratory research, achieving perfect real-world scenarios is nearly impossible. Most obviously, students doing homework often need to learn materials that are highly rich and complex, and also diverse. For instance, materials typically include math, science, reading, history, and a host of other subjects. Thus, the strategies that people use to allocate, order, or space their study might be influenced by a myriad of untested factors. However, in the real world, very simple and comparable materials are also studied and learned by people of all ages. Vocabulary learning, both of advance vocabulary in one's native language and in foreign language learning, is a crucial skill necessary starting at young ages and continuing into high-school, college, and beyond. From this perspective, the experimental materials used here are certainly appropriate for generalizing to crucial aspects of real-world learning.

Although homework is sometimes carried out in groups, it is typically a personal and private practice. When a student does well in school, both parents and teachers praise that student alone; when a student does poorly in school, parents and teachers are concerned about that individual's studying strategies. Because of the notion that study consists of personal decisions, individual judgments of learning and one's control of strategies are vital. Thus, in empirical research, assessing an item's difficulty by taking individual JOLs seems to be a realistic measure of learning. As a result, understanding how these judgments can guide the various decisions students make is crucial.

In order to get a richer idea of why students make the decisions they do, it might be beneficial to expand such metacognitive procedures to include open-ended or longer verbal protocols. For instance, it remains unknown whether our learners were aware of the benefits of using spacing strategies rather than massing strategies. There is an advantage of examining how people *actually* choose to study, as we do here, as opposed to verbal reports of how they *say* they study, however: People may simply be unaware of what strategies they use and why they use them. Verbal protocols and behavioral data can both contribute to the study of metacognition and learning, their relationship is an important question for future research, both in the classroom and in the laboratory.

In short, we believe that the methods used here represent a significant advance in the degree of ecological validity of research on study decisions as they related to time allocation, ordering, and spacing. Moreover, combining the three types of decision had benefits beyond ecological validity; by asking participants to make all three decisions in concert, we gained insight into how people jointly manage these three decisions, and how each decision relates to metacognitive judgments. This allowed us to determine, for example, that people's individual metacognitive knowledge guides the choices that they make, in particular, those of time allocation and or the order in which they study the various items.

Conclusion

To maintain experimental control, much of the research on study decisions has focused on one particular decision (e.g. time allocation). Here, while maintaining control, we investigated three decisions in concert: time allocation, ordering, and spacing. Making such complex decisions may be one of the most challenging experiences learners face, especially given that many students report having received little or no instruction on how to study effectively (Kornell & Bjork, 2007). Nevertheless, the data on metacognitive control presented here indicate that our participants met such challenges more than adequately.

References

Benjamin, A. S. & Bird, R. (2006). Metacognitive control of the spacing of study repetitions. *Journal of Memory and Language, 55*, 126-137.

Cepeda, N J, Pashler, H, Vul, E, Wixted, J, & Rohrer, D (2006). Distributed practice in verbal recall tasks: A review and quantitative synthesis. *Psychological Bulletin*, *132*, 354-380.

Dempster, F. N. (1989). Spacing effects and their implications for theory and practice. *Educational Psychology Review, 1*, 309-330.

Dunlosky, J., & Hertzog, C. (1998). Training programs to improve learning in later adulthood: Helping older adults educate themselves. In D. J. Hacker, J. Dunlosky, & A. Graesser (Eds.), *Metacognition in educational theory and practice* (pp. 249–276). Hillsdale, NJ: Lawrence Erlbaum.

Gilovich, T., Vallone, R. & Tversky, A. (1985). The hot hand in basketball: On the misperception of random sequences. *Cognitive Psychology, 17*, 295-314.

Glenberg, A. M. (1979). Component-levels theory of the effects of spacing of repetitions on recall and recognition. *Memory & Cognition, 7*, 95-112.

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

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

Kornell, N., & Metcalfe, J. (2006). Study efficacy and the region of proximal learning framework. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 32,* 609-622.

Kornell, N., & Son, L. K. (in press). Learners' choices and beliefs about self-testing. *Memory.*

Metcalfe, J., & Kornell, N. (2003). The dynamics of learning and allocation of study time to a region of proximal learning. *Journal of Experimental Psychology: General, 132,* 530-542.

Metcalfe, J., & Kornell, N. (2005). A region of proximal learning model of study time allocation. *Journal of Memory and Language, 52*, 463– 477.

Nelson, T. O., & Leonesio, R. J. (1988). Allocation of self-paced study time and the"labor-in-vain effect." *Journal of Experimental Psychology: Learning, Memory,and Cognition, 14,* 676–686.

Nelson, T. O., & Narens, L. (1990). Metamemory: A theoretical framework and new findings. In G. H. Bower (Ed.), *The psychology of learning and motivation* (Vol. 26, pp. 125–173). New York: Academic Press.

Nelson, T. O., Dunlosky, J., Graf, A., & Narens, L. (1994). Utilization of metacognitive judgments in the allocation of study during multitrial learning. *Psychological Science, 5,* 207-213.

Simon, D. A., & Bjork, R. A. (2001). Metacognition in motor learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 27*, 907-912.

Son, L. K. (2004). Spacing one's study: Evidence for a metacognitive control strategy. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 30*, 601-604.

Son, L. K., & Kornell, N. (2008). Research on the allocation of study time: Key studies from 1890 to the present (and beyond). In J. Dunlosky & R. A. Bjork (Eds.), *Handbook of metamemory and memory* (pp. 333–351). New York: Taylor & Francis.

Thiede, K. W., & Dunlosky, J. (1999). Toward a general model of self-regulated study: An analysis of selection of items for study and self-paced study time. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 25*, 1024–1037.

Zechmeister, E. B., & Shaughnessy, J. J. (1980). When you know that you know and when you think that you know but you don't. *Bulletin of the Psychonomic Society, 15*, 41-44.

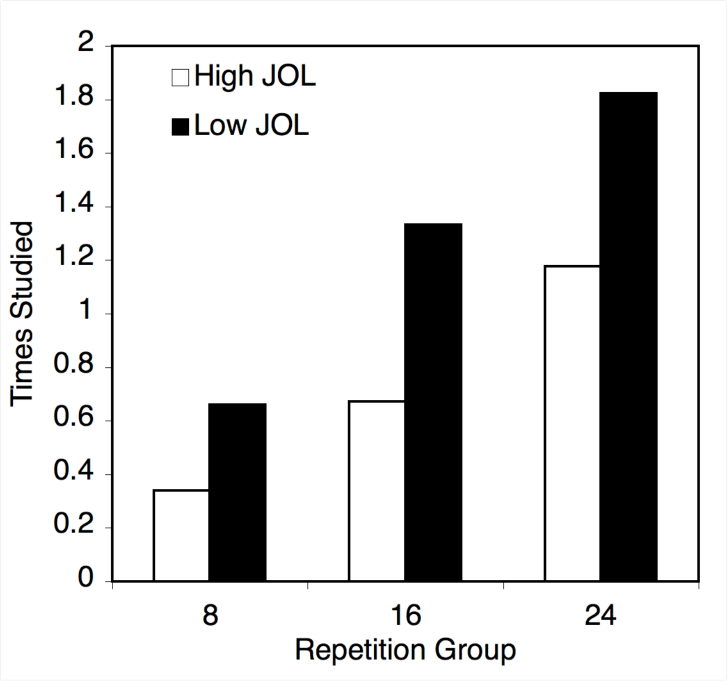
Acknowledgements

This research was supported by the Spencer Foundation. Correspondence concerning this article should be addressed to Lisa Son, Department of Psychology, Barnard College, 3009 Broadway, New York, NY 10027. Electronic mail may be sent to lson@barnard.edu or nkornell@ucla.edu.

Figure Captions

Figure 1. Number of times items were studied as a function of JOL (categorized into High or Low) and repetition group.

Figure 2. Random, actual, and maximum possible spacing.



2
1.8
1.6
1.4
1.2
1
0.8
0.6
0.4
0.2
0
Times Studied
□ High JOL
■ Low JOL
8
16
24
Repetition Group

Figure 1

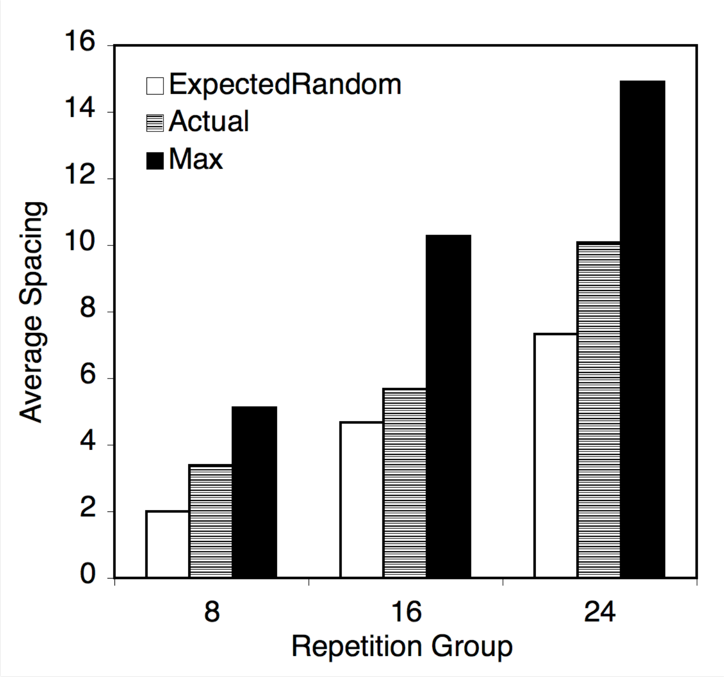

16
14
12
10
8
6
4
2
0
Average Spacing
ExpectedRandom
Actual
Max
8
16
24
Repetition Group

Figure 2