Metacognitive Control Over the Distribution of Practice: When Is Spacing Preferred?

Thomas C. Toppino, Michael S. Cohen, Meghan L. Davis, and Amy C. Moors
Villanova University

The authors clarify the source of a conflict between previous findings related to metacognitive control over the distribution of practice. In a study by L. Son (2004), learners were initially presented pairs of Graduate Record Examination (GRE) vocabulary words and their common synonyms for 1 s, after which they chose to study the pair again immediately (massed practice), later (spaced practice), or not at all (done). Learners chose spaced practice less as pair difficulty increased. A. S. Benjamin and R. D. Bird (2006), using different materials and procedures and a longer presentation duration (5 s), concluded just the opposite. The authors adopted Son's materials and procedures and replicated her findings with a 1-s stimulus duration. However, the declining choice of spacing as item difficulty increased largely reflected learners' failure to fully perceive items with brief presentations. With longer presentations, ensuring full perception, the choice of spaced practice increased with greater pair difficulty, in agreement with Benjamin and Bird. Theoretical implications are discussed in the context of discrepancy-reduction and proximal-learning perspectives.

Keywords: metacognition, spacing effect, metacognitive control, distributed practice

The study of learning and memory historically has focused on how the environment shapes our knowledge and behavior, but it has become increasingly clear that critical roles are played by people's strategies and by their awareness of their own knowledge and cognitive processes (metacognition). People often are relatively accurate in their ability to monitor their cognitive processes. This is demonstrated, for example, by the fact that judgments of learning (JOLs) made after one studies target material usually correlate positively with objective assessments of learning. People also use their metacognitive knowledge to control their study strategies in presumably adaptive ways.

A sizeable literature on the metacognitive control of learning strategies has focused on how learners allocate study time (or opportunities) to the items they are trying to acquire. A dominant theory in this area has been the *discrepancy reduction* model (Dunlosky & Hertzog, 1998), according to which learners adopt a target level of mastery and preferentially allocate resources to items for which the judged level of learning is most discrepant from the desired level. Thus, learners should devote more study time to items that they judge to be more difficult and less well learned. Another major theory, the *region of proximal learning*

model (Metcalfe, 2002; Metcalfe & Kornell, 2005), leads to different predictions when learning resources are limited. For example, when study time is restricted, learners should be better off devoting their limited study time to moderately difficult or moderately well-learned items rather than to the most difficult, least-learned items. A key assumption is that the decision to continue studying or not depends on metacognitive judgments of the rate of learning (jROLs), which reflect the rate of information uptake. Thus, people should allocate more study time to items for which jROLs are high, such as moderately difficult items that are not yet fully learned. People should be less likely to continue studying items with lower jROLs, including the easiest items that are already learned and the hardest items on which little progress can be expected.

Both the discrepancy reduction and proximal learning approaches have amassed considerable empirical support from studies examining learners' choices with respect to which items should be given additional study time (e.g., Dunlosky & Hertzog, 1997; Dunlosky & Thiede, 1998; Metcalfe & Kornell, 2003, 2005; Nelson & Leonesio, 1988; Son & Metcalfe, 2000). However, it is also possible to ask how learners would distribute their practice over time when the total amount of study time per item is held constant. There has been a massive amount of research indicating that spaced or distributed study leads to better retention than massed study (for reviews, see Cepeda, Pashler, Vul, Wixted, & Rohrer, 2006; Crowder, 1976; Dempster, 1996; Hintzman, 1974). However, there is relatively little research on people's metacognitive awareness of the benefits of spaced practice and even less on how learners will distribute practice when they are allowed to determine this for themselves.

When metacognitive monitoring is assessed, learners often fail to appreciate the benefits of spaced practice relative to massed practice (e.g., Baddeley & Longman, 1978). In the context of list

learning, participants give spaced items higher JOLs than massed items when judgments are delayed (Dunlosky & Nelson, 1994) but not when judgments are made immediately after a repeated item's second presentation (Zechmeister & Shaughnessy, 1980). Immediate JOLs apparently are vulnerable to transient cues (e.g., retrieval fluency) that cause people to overestimate the benefits of massed practice (Benjamin, Bjork, & Schwartz, 1998).

With respect to metacognitive control of the distribution of practice, very little research has been conducted, and the results have been contradictory. Son (2004) introduced the topic with a study in which participants were asked to learn pairs consisting of Graduate Record Examination (GRE) vocabulary words and their common synonyms. A pair was presented initially for 1 s. Then, learners made a JOL and were prompted to make one of three choices. They could choose not to study the item again so that they were "done" with it. They could choose to study it again "now," in which case the item was presented again immediately (massed practice). Or, they could choose to study it again "later" after all other pairs had been presented at least once (spaced practice). Following list presentation and a distractor task, a cued recall test was administered.

Son's (2004) results indicated that the proportion of re-studied items selected for spaced practice declined as JOLs got lower. In other words, participants tended to prefer spaced practice for easy items, but their tendency to choose spaced practice declined and their tendency to choose massed practice increased as items became more difficult. Cued recall was better for spaced items regardless of the level of JOL, although systematic analyses of the recall data were impossible because spacing was confounded with item-selection and retention-interval effects that were inherent in Son's procedure.

Son (2004) interpreted her critical choice data in terms of a metacognitive strategy in which learners choose to continue studying (i.e., select massed practice) when their knowledge is low, but choose to delay re-studying (i.e., opt for spaced practice) when their knowledge is relatively high. Although she did not explicitly make the connection, Son's results seem to be consistent with the region of proximal learning model. That is, after studying hard items for 1 s, jROLs may still be high so that further immediate study would seem fruitful. However, after easy items are studied for 1 s, jROLs may indicate that the rate of information uptake is already slowing so that further study would be more efficient if it were postponed.

Benjamin and Bird (2006) investigated the issue further using a different set of procedures. In their first two experiments, they varied the difficulty of pairs of common words and presented all pairs twice, for 5 s on each occurrence. Following the first presentation, participants made a JOL (in Experiment 1 only) and then chose to study the item again either "sooner" (massed practice) or "later" (spaced practice). If they selected "sooner," the pair's second presentation occurred after the presentation of one other pair. If they selected "later," the pair was re-presented after all other pairs had been presented at least once. Participants' selections were restricted so that they had to choose spaced practice for exactly one-half of the items. Finally, a cued recall test was administered.

A conceptual replication of Son's (2004) findings would have been achieved if Benjamin and Bird (2006) had found that participants allocated more of their limited spaced-practice trials to easy

items and more of their massed-practice trials to difficult items. However, they obtained just the opposite result in Experiments 1 and 2. Participants chose spacing more often for difficult items and massing more often for easy items. In a third experiment, Benjamin and Bird used extremely fast presentation times (0.5 s for each presentation of a pair with a 0.1-s interpair interval) and found no difference in learners' tendency to space repetitions for easy and difficult items. They concluded that their results were most consistent with discrepancy reduction, according to which the more efficient study strategy (spacing) generally should be preferred for all unlearned items and, if anything, might be preferentially allocated to the most difficult items (i.e., those most discrepant from the target level of learning).

The conflicting findings obtained by Son (2004) and by Benjamin and Bird (2006) preclude any general conclusions about when and why learners choose to mass or space practice. In the present experiments, we hoped to learn how the contradictory findings might be reconciled and, more generally, we sought to clarify the metacognitively controlled strategies that learners use in deciding how to distribute practice.

Experiment 1

The most obvious methodological difference between studies by Son (2004) and by Benjamin and Bird (2006) was that Son's procedure placed no limits on how many items could be massed or spaced, whereas Benjamin and Bird required that half of the items receive each kind of repetition. Another salient difference was that Son, unlike Benjamin and Bird, gave learners the option of being "done" with an item rather than re-studying it. She included the "done" option because previous work (Son, 2002) had indicated that without it, some learners would choose to mass some of the easiest (already learned) items just to get them out of the way. Implementing a "done" option in the context of Benjamin and Bird's procedure would have been challenging because of the difficulty of equating the number of massed and spaced choices when the number of items chosen for re-studying is under participants' control. However, the absence of a "done" option could have contributed to the discrepant results if some of Benjamin and Bird's participants adopted the strategy of massing easy items in order to dispense with them.

A third potentially important difference was that the presentation duration of pairs was much longer in Benjamin and Bird's (2006) study (5 s for both presentations) than in Son's (2004) study (1 s for the first presentation and 3 s for the second). From the region of proximal learning perspective, longer durations might increase the preference for spacing because the perceived rate of learning (jROLs) may decline for difficult as well as for easy items over a long study period, leading to a general preference for spacing. Furthermore, there is some empirical evidence that timing can be a critical variable in this context. In contrast to her 2004 data, Son (2002) obtained a general preference for spaced practice and no consistent effect of item difficulty when she used a longer initial presentation duration (3 s). And Benjamin and Bird (Experiment 3) found that a very fast (0.5 s) presentation duration eliminated the preference for spacing difficult items that they had obtained with a longer presentation duration (5 s).

In Experiment 1, we systematically explored the contribution of timing. The initial presentation duration of pairs was either 1 s,

corresponding to Son's (2004) brief duration, or 5 s, corresponding to Benjamin and Bird's (2006) comparatively long duration. In general, we adopted Son's (2004) methodology because it allows learners to use the "done" option for items that they judge to be already learned. However, we differed from Son by not soliciting JOLs from our participants. Instead, we varied item difficulty experimentally. This controls complex Participant × Item interactions that could have contributed in unspecified ways to Son's results. Additionally, previous research has shown a close correlation between JOLs and a priori measures of difficulty (e.g., Benjamin & Bird, 2006).

Method

Participants and design.   The experimental design was a 2 × 3 mixed factorial involving two presentation durations (1 s vs. 5 s) and three levels of item difficulty (easy, medium, and hard). The latter factor was manipulated within participants.

Ninety-one introductory psychology students participated for course credit. They were assigned randomly to the two between-participants presentation-duration conditions such that 1-s and 5-s groups consisted of 46 and 45 participants, respectively. The data from an additional 6 participants were eliminated. Two failed to re-study any items from at least one item-difficulty level, and 4 were excluded as a result of computer failure.

Materials.   Participants studied 48 cue–target pairs, consisting of GRE vocabulary words (cues) and their more common synonyms (targets). They were selected from the 60 pairs used by Son (2004) who also graciously provided the JOL data from her 2004 study for each pair. We used the JOL data as an index of item difficulty to select 16 pairs in each of three ranges of judged difficulty (easy, medium, and hard).

Procedure.   Before list presentation, participants were thoroughly instructed about the procedure and about each strategy choice they would have the opportunity to make. They further were instructed that they should choose the strategy that they believed would help them the most in remembering the target in a later cued recall test.

During list presentation, intact pairs (cue and target together) were presented successively on a computer screen in an order that was randomized independently for each participant. After a pair's initial presentation, for either 1 s or 5 s depending on the condition, participants were given a choice between "study now" (massed practice), "study later" (spaced practice), and "done" (no additional practice). If "done" was chosen, the word pair was not shown again. If massed practice was chosen, the pair was shown again immediately for 3 s, following Son (2004). If spaced practice was chosen, the pair was added to a queue and, following the procedures of both Son (2004) and Benjamin and Bird (2006), was re-presented for 3 s after all items had been presented at least once. Thus, the second occurrences of spaced items were presented at the end of the list in a new random order.

After a 5 min distractor task in which participants solved simple arithmetic problems, a cued recall test was administered. Cue (GRE) words were presented in a new random order. Participants were given 10 s to enter the target into the computer before the next cue was presented.

Results

Significance was set at $p = .05$ for all analyses in this article unless otherwise specified.

If our a priori manipulation of item difficulty was successful, participants should become less likely to choose "done" (no additional studying) as item difficulty increases. In agreement with this prediction, a 2 × 3 (Presentation Duration × Item Difficulty) analysis of variance (ANOVA) with repeated measures on the second factor indicated that the proportion of "done" responses declined significantly as item difficulty increased, $F(2, 178) = 89.08$, $MSE = .012$ (see the top panel of Figure 1). No other effects were reliable ($Fs \leq 1.00$).

The question of primary interest was how presentation duration and item difficulty affected participants' choice of spacing or massing practice when items were re-studied. The proportion of re-studied items for which spaced practice was selected (see Figure 2) was submitted to a 2 × 3 (Presentation Duration × Item Difficulty) ANOVA with repeated measures on the second factor. There was a main effect of presentation duration such that a greater proportion of re-studied items were spaced in the 5-s than in the 1-s condition, $F(1, 89) = 11.72$, $MSE = .25$. Although there was no main effect of item difficulty, $F(2, 178) < 1.00$, the two-way interaction was reliable, $F(2, 178) = 9.63$, $MSE = .02$. As item difficulty increased, the proportion of re-studied items that were selected for spaced practice significantly decreased in the 1-s condition, $F(2, 90) = 3.47$, $MSE = .02$, and significantly increased in the 5-s condition, $F(2, 88) = 6.50$, $MSE = .02$.

Cued recall performance was of minor interest in the present experiment. As in Son's (2004) study, it was completely confounded by complex Participant × Item interactions such that no statistical analysis seemed entirely appropriate. Nevertheless, two points may be of interest. First, recall was greater for spaced items (29.47%) than for massed items (13.94%). Second, recall declined with increasing item difficulty (42.43%, 23.06%, and 13.75% for easy, medium, and hard items, respectively). These findings provide added confirmation that our manipulation of item difficulty was effective.

Discussion

When the first presentation of each pair was brief, we replicated Son's (2004) finding that the proportion of spaced items declined with increasing difficulty. In contrast, when we used a relatively long duration, we found that the proportion of spaced items increased as a function of item difficulty, which is consistent with the results reported by Benjamin and Bird (2006). These results suggest that the discrepant findings obtained by Son and by Benjamin and Bird are most likely attributable to a variation in timing rather than to other differences between the studies. However, it remains unclear why item difficulty affects the choice of spaced practice so differently in the two timing conditions. Neither Son's hypothesis, which is related to a proximal learning account, nor discrepancy reduction seem capable of explaining both outcomes. The next experiment explored the possibility that one of these outcomes was an unintended artifact of the procedures that were used.
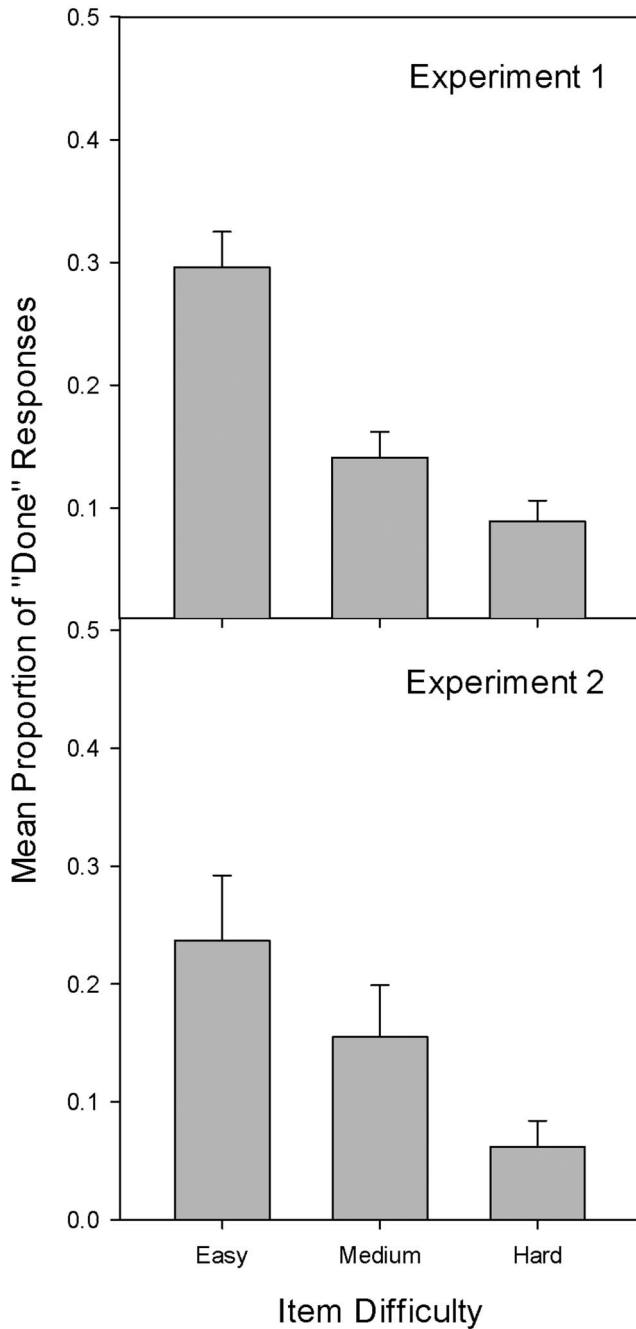
Figure 1. The mean proportion of "done" responses selected as a function of item difficulty in Experiments 1 and 2.

Experiment 2

A number of participants in the 1-s condition of Experiment 1 reported that the initial stimulus presentation had been so fast they had been unable to read both members of some pairs. We became concerned that participants were having difficulty perceiving the full stimulus pair and establishing an initial encoding (i.e., a stable percept). Moreover, this problem likely would be exacerbated by the more difficult pairs, which included longer, less familiar words

than the easier pairs. A very reasonable strategy if one fails to encode a stimulus would be to ask to see it again right away (massed presentation). Participants simply might want to know what it is that they just failed to perceive. Or, they might intuit that they would be more likely to perceive the item if it was re-presented immediately, in which case they might benefit somewhat from whatever partial perception was achieved on the first presentation. In either case, failure to perceive would lead to more massing and less spacing as difficulty increased.

In this experiment, we replicated the 1-s condition of Experiment 1, except for a single modification. Immediately after participants made their choice with respect to re-studying an item, they were prompted to indicate whether or not they had seen the entire pair. We were interested in whether failures to perceive the full stimulus increased with item difficulty and the extent to which that could account for the 1-s results of Experiment 1.

Method

Participants and design. The experiment involved a one-way, within-participants design in which the single independent variable was item difficulty (easy vs. medium vs. hard). Twenty-nine introductory psychology students participated for course credit. Data from an additional 3 students were excluded because they either chose not to re-study any items in at least one difficulty level or failed to follow instructions.

Materials and procedure. The materials and procedures were identical to the 1-s condition of Experiment 1, with one exception. After choosing a study strategy for each item, participants were required to respond (yes/no) to a prompt asking whether they had seen both words of the pair.

Results and Discussion

A one-way, repeated-measures ANOVA on the proportion of "done" responses indicated that participants became less likely to choose "done" as item difficulty increased, $F(2, 56) = 10.76$,
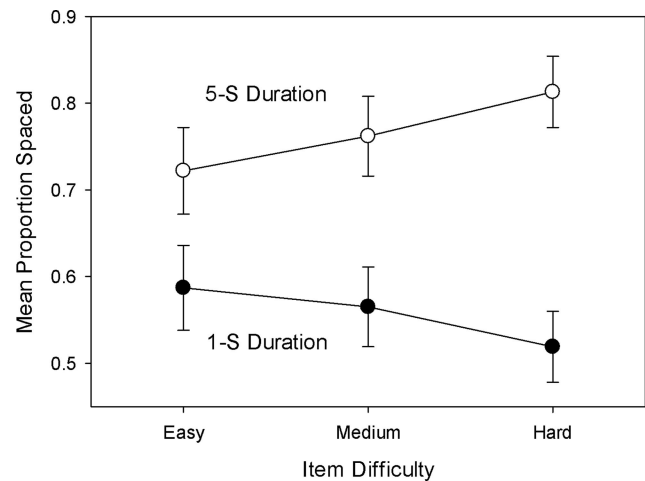

Figure 2. The mean proportion of re-studied items that were selected for spaced practice in Experiment 1 as a function of initial presentation duration and item difficulty.

MSE = .021, replicating the results of Experiment 1 (see the lower panel of Figure 1).

The results of primary interest were explored with three one-way, repeated-measures ANOVAs. First, the mean proportion of successfully perceived items declined significantly with increasing item difficulty, $F(2, 56) = 37.03$, $MSE = .011$ (see Figure 3). Thus, as suspected, more difficult pairs were less likely to be fully perceived. Second, when the proportion of pairs that were selected for spaced practice was based on all pairs (perceived or not), the proportion of spaced pairs declined significantly as item difficulty increased, as shown in the lower curve of Figure 4, $F(2, 56) = 3.91$, $MSE = .016$. This confirms the replicability of the results of Son's (2004) study and of our Experiment 1, while providing reassurance that querying participants about whether they had fully seen each pair did not markedly alter the results. Third, we restricted the data to pairs that participants claimed to have fully perceived, and we re-examined the effect of item difficulty on the proportion of items that were selected for spacing. As indicated in the upper curve of Figure 4, the effect of item difficulty was greatly reduced and did not approach significance, $F(2, 52) < 1.00$, $MSE = .015$.[1]

These results suggest that the primary factor accounting for the decrease in the choice of spacing as item difficulty increased is that more difficult items were less likely to be fully perceived. Participants chose to mass 66.6% of the items that they reported not perceiving but less than 30% of those that they saw. If anything, the influence of the failure to perceive the items fully is likely to have been underestimated in our results. We asked learners whether they had seen both members of a pair because this was easy to communicate, but it may be possible to "see" both members of a pair in some literal sense without being able to fully encode them both by establishing a stable representation. Thus, it seems likely that some pairs were identified as having been seen even though they were not fully encoded. Nevertheless, our findings suggest that much, if not all, of the effect of item difficulty that was obtained in the 1-s duration condition of Experiment 1 and in Son's (2004) study was an artifact of not being able to fully perceive rapidly presented pairs involving relatively rare words. When a pair is not fully perceived, participants have a strong tendency to ask to see it again immediately (massed presentation).
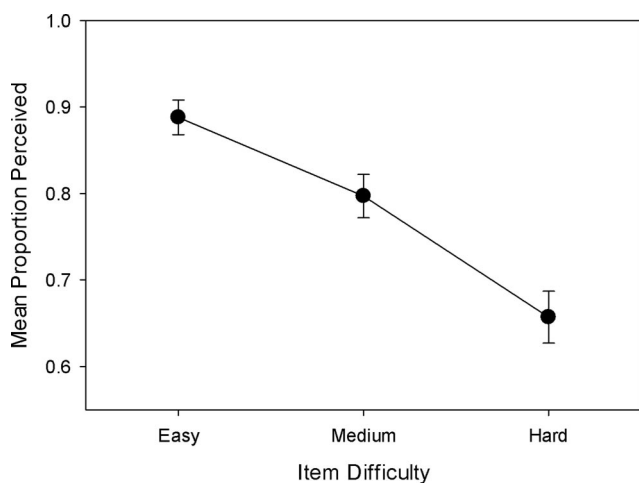

Figure 3. The mean proportion of items that were reported to have been fully perceived in Experiment 2 as a function of item difficulty.
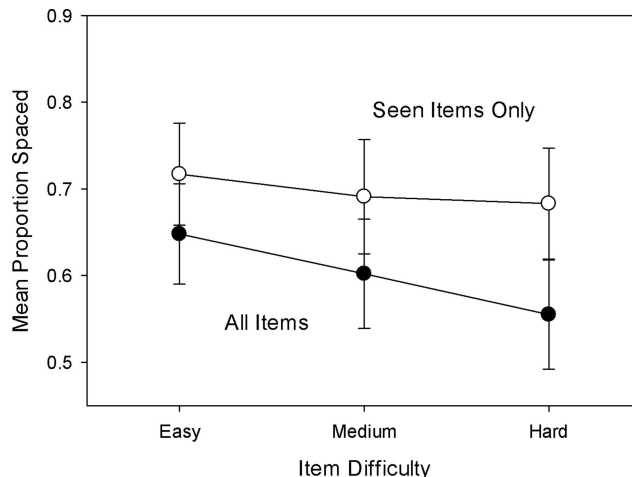

Figure 4. The mean proportion of re-studied items that were selected for spaced practice in Experiment 2 as a function of spaced practice. The upper curve represents the data when only perceived items were included. The lower curve shows the data when all items were included in the analysis.

Finally, the cued recall results showed the same pattern as in Experiment 1. The percentage of targets that were remembered in the cued recall test was higher for items that had been spaced (27.86%) than for items that had been massed (15.90%). Also, the percentage of targets recalled decreased as item difficulty increased (43.08%, 20.31%, and 8.26% for easy, medium, and hard items, respectively).

General Discussion

The present research clarified the source of a major discrepancy in the literature concerning the strategies that learners use in deciding whether practice should be massed or distributed. Son (2004) found that learners' preference for spaced practice was greatest for the easiest items and declined with increasing item difficulty. Benjamin and Bird (2006) found that learners were more likely to choose distributed practice for more difficult items. Although these previous studies differed in numerous ways, we found in Experiment 1 that the duration of the initial presentation of to-be-learned pairs played a crucial role in accounting for the discrepant results. When we used a short initial-presentation duration of 1 s, as Son did, we replicated her findings. The tendency to choose spaced practice declined as item difficulty increased. However, when we used a relatively long presentation duration of 5 s, as Benjamin and Bird did, our results were consistent with theirs. Learners chose spaced practice more often as items became more difficult.

The results of Experiment 2 provided insight into why the choice of spaced practice declined and the choice of massed practice increased with item difficulty when a short 1-s presenta-

---

[1] Two participants whose data were included in the analysis based on all re-studied pairs were excluded from the conditionalized analysis because, in at least one difficulty condition, they did not have any items that were both seen and chosen for re-study. It should be noted that removing these participants' data from the unconditionalized analysis based on all items regardless of whether they were perceived did not change the results of that analysis.

tion interval was used. Learners had difficulty encoding a stable perceptual representation of the items comprising a pair when presentations were very brief. This problem became greater as item difficulty increased, presumably because the GRE vocabulary cues for more difficult pairs consisted of longer and rarer words. When learners did not fully perceive a pair, they exhibited a strong preference for seeing the item again right away (massed practice). They may implicitly believe that an item will be easier to perceive if it occurs again immediately while a trace of the previous partial perception is still available. Thus, choosing to see an item immediately after failing to fully perceive it may reflect a metacognitive process. Our point, however, is that the decision seems to be more related to perceptual processing than to learning.

We originally had conceptualized the reduction in spacing that Son (2004) had found as a function of increasing item difficulty to be consistent with predictions derived from the region of proximal learning model (e.g., Metcalfe, 2002; Metcalfe & Kornell, 2005). A crucial assumption of this model is that learners will continue studying when jROLs are high but stop when jROLs are low. The latter could occur either because learning has been largely accomplished (at least for the moment) or because the items are so hard that little progress is being made. In the context of the present study, this implies that, following a brief study period, jROLs will be low for easy items because people will have learned about as much as they can and that jROLs will be higher for more difficult items (Metcalfe & Kornell, 2005). Thus, the choice of spacing should decline and massing should increase as item difficulty becomes greater. Although this pattern of results was obtained by Son (2004) and by us in Experiment 1, it apparently was obtained for different reasons than those hypothesized by proximal learning. The results of Experiment 2 suggest that the choice of massing difficult items was not a consequence of high jROLs but instead was largely the result of perceptual failures that would be expected to be associated with low jROLs.

On the basis of the perceptual failure interpretation of our results, we would not expect brief presentations to lead to the same effect under conditions in which the difficulty of learning items is uncorrelated with the difficulty of perceiving them. Some support for this contention was found in a recent study by Cohen (2007) using procedures similar to those used by us and by Son (2004). Cohen avoided the correlation between item difficulty and perceptual difficulty by using pairs of common words while varying the difficulty of learning the pairs by means of the normative relationship between pair members. Several aspects of the results (e.g., the frequency of choosing "done" responses) provided strong evidence that this method of varying item difficulty was effective, but there was no hint of a decline in the choice of spacing as item difficulty increased even though a brief, 1-s exposure duration was used. In fact, there was a pronounced, though nonsignificant, trend for spacing to be chosen more often for harder items.

The use of longer presentation durations can be construed as another way of ensuring item perception. When we used a long duration in Experiment 1, the tendency for learners to choose spacing less for more difficult items was not only eliminated, it was reversed. Thus, our results can be summarized as indicating that, when perceptual difficulties are avoided, learners prefer spaced practice over massed practice, and their preference for spacing becomes greater as item difficulty increases. The remaining question is why.

From a proximal learning perspective, it is possible that, after a long initial presentation duration (e.g., 5 s), learning may be largely finished (at least temporarily) for items at all difficulty levels, leading to low jROLs across the board. In this case, one might expect that there would be a strong preference for spacing and that this preference would be similar regardless of item difficulty. It is unclear whether the theory can be adapted to account for the fact that learners exhibited a greater tendency to space difficult items than easy items.

A discrepancy reduction-based account may be more promising. Benjamin and Bird (2006) suggested that a preference for spacing more difficult items may be generally consistent with discrepancy reduction because learners allocate the more effective strategy to the harder items that need it most. This can be stated most simply by proposing that the attractiveness of spaced practice is proportional to the discrepancy between the degree to which an item is currently learned and the normative level to which the learner aspires. This hypothesis is consistent with Benjamin and Bird's results and with our findings when we used a long 5-s presentation duration. However, it begs the question of why learners, who clearly appreciate the value of spaced practice, do not maximize their gains by choosing to space all items when procedures allow it as was the case in our experiments.

One possibility is that learners may perceive there to be some kind of cost associated with distributed practice relative to massed practice. In this case, the observed pattern of results might be explained in terms of learners choosing spaced practice for items that need it most while avoiding the costs of spaced practice when its benefits for learning are needed less. Although such a cost might be conceptualized in a variety of ways, an illustrative analysis can be offered in terms of the costs and benefits of processing effort (i.e., the expenditure of cognitive resources). There is evidence that processing the second occurrence of a re-studied item is more effortful when the item receives spaced as opposed to massed practice (e.g., Johnston & Uhl, 1976; Magliero, 1983). The assumed benefit of greater processing effort underlies many accounts of the spacing effect, including most deficient processing theories (Dempster, 1996) and some study-phase retrieval theories (e.g., Thios & D'Agostino, 1976). However, expending cognitive resources usually comes at a cost. If learners have difficulty sustaining a high level of cognitive effort over an extended period of time, such as that required for the presentation of a list of to-be-remembered items as suggested by Greeno (1970), they might conserve resources on some items by massing practice so that they can exert more effort on other items by spacing practice. The most adaptive strategy might be to mass unlearned items that do not require much additional study (i.e., easier items) and to space harder items that are in the greatest need of more effortful processing.[2]

In closing, our research clarified the basis of a discrepancy in the literature related to learners' metacognitive decisions about how to

---

[2] A variation on the processing effort hypothesis is that sustaining a high level of processing effort may be aversive (perhaps tiring) in a way that is analogous to the effect of intense physical or motor effort. Thus, massing some items may be a way of reducing or avoiding the aversive consequences of spacing all items. We note that this analysis is entirely compatible with the more cognitive analysis provided in the main text.

distribute their practice during learning. When perceptual difficulties are avoided, we found that learners prefer distributed practice over massed practice and that this preference becomes greater as item difficulty increases. Further research is needed, however, to determine why the preference for spaced practice increases with item difficulty. For example, is there a cost to engaging in spaced practice that leads learners to allocate it selectively to harder items?

References

Baddeley, A. D., & Longman, D. J. A. (1978). The influence of length and frequency of training session on the rate of learning to type. *Ergonomics, 21,* 627–635.

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

Benjamin, A. S., Bjork, R. A., & Schwartz, B. L. (1998). The mismeasure of memory: When retrieval fluency is misleading as a metamnemonic index. *Journal of Experimental Psychology: General, 127,* 55–68.

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,* 354–380.

Cohen, M. S. (2007). *How do subjects use judgments of item difficulty to guide study strategies in selection of spaced or massed practice? A comparison of theories.* Unpublished master's thesis, Villanova University.

Crowder, R. G. (1976). *Principles of learning and memory.* Hillsdale, NJ: Erlbaum.

Dempster, F. (1996). Distributing and managing the conditions of encoding and practice. In E. L. Bjork & R. A. Bjork (Eds.), *Memory* (pp. 317–344). San Diego: Academic Press.

Dunlosky, J., & Hertzog, C. (1997). Older and younger adults use a functionally identical algorithm to select items for restudy during multitrial learning. *Journal of Gerontology: Psychological Sciences, 52,* 178–186.

Dunlosky, J., & Hertzog, C. (1998). Training programs to improve leaning in later adulthood: Helping older adults educate themselves. In D. J. Hacker, J. Dunlosky, & A. C. Graesser (Eds.), *Metacognition in educational theory and practice* (pp. 249–275). Mahwah, NJ: Erlbaum.

Dunlosky, J., & Nelson, T. O. (1994). Does the sensitivity of judgments of learning (JOLs) to the effects of various study activities depend on when the JOLs occur? *Journal of Memory and Language, 33,* 545–565.

Dunlosky, J., & Thiede, K. W. (1998). What makes people study more? An evaluation of factors that affect people's self-paced study. *Acta Psychologica, 98,* 37–56.

Greeno, J. G. (1970). Conservation of information-processing capacity in paired-associate memorizing. *Journal of Verbal Learning and Verbal Behavior, 9,* 581–586.

Hintzman, D. L. (1974). Theoretical implications of the spacing effect. In R. L. Solso (Ed.), *Theories in cognitive psychology: The Loyola Symposium* (pp. 77–99). Hillsdale, NJ: Erlbaum.

Johnston, W. A., & Uhl, C. N. (1976). The contributions of encoding effort and variability to the spacing effect on free recall. *Journal of Experimental Psychology: Human Learning and Memory, 2,* 153–160.

Magliero, A. (1983). Pupil dilations following pairs of identical words and related to-be-remembered words. *Memory & Cognition, 11,* 609–615.

Metcalfe, J. (2002). Is study time allocated selectively to a region of proximal learning? *Journal of Experimental Psychology: General, 131,* 349–363.

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.

Son, L. K. (2002). *Spacing one's study: Evidence for a metacognitive control strategy.* Paper presented at the annual meeting of the Psychonomic Society, Kansas City, MO.

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., & Metcalfe, J. (2000). Metacognitive and control strategies in study-time allocation. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 26,* 204–221.

Thios, S. J., & D'Agostino, P. R. (1976). Effects of repetition as a function of study-phase retrieval. *Journal of Verbal Learning and Verbal Behavior, 15,* 529–536.

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.