

Toward an understanding of students' allocation of study time: Why do they decide to mass or space their practice?

MARY A. PYC AND JOHN DUNLOSKY
Kent State University, Kent, Ohio

Two recent studies have led to discrepant findings regarding students' decisions to mass or space practice of to-be-learned items: Son (2004) reported that participants chose to mass practice of difficult items more than that of easy items, whereas Benjamin and Bird (2006) reported the opposite. To resolve this apparent discrepancy, we had participants study items during an initial trial and then decide whether the next study trial would be massed or spaced. Across three experiments, we systematically varied factors that differed across these studies. In general, the participants more often chose to space than to mass practice, but they did favor massed practice when items were very difficult to learn. Moreover, although previous hypotheses implicated metamemory causes for these effects, the present results indicate that nonmetamemory causes are also responsible, such as using spaced practice as a default and deciding to mass practice for items that had not been fully perceived during an initial study trial.

The influence of massed versus spaced practice on acquisition and retention has been extensively investigated since Ebbinghaus (1885/1913; for reviews, see Cepeda, Pashler, Vul, Wixted, & Rohrer, 2006; Dempster, 1988; Donovan & Radosevich, 1999; Hintzman, 1974; Melton, 1970). Both massed practice and spaced practice involve multiple study trials with to-be-learned items. For massed practice, all study trials for a given item occur consecutively, whereas for spaced practice, trials for a given item are distributed across time, which is usually filled with the practice of other to-be-learned material. Although research has firmly established that retention is usually far superior after spaced than after massed practice, relatively little attention has been paid to how students allocate practice across to-be-learned materials when they control learning themselves. If they do not spontaneously space practice and instead overrely on massed practice, a major goal of research would include discovering how best to train and motivate them to use this more effective practice schedule.

Son (2004) and Benjamin and Bird (2006) evaluated how students make allocation decisions regarding whether to mass or space practice of to-be-learned information, and a major result from them appears inconsistent. Son reported that participants more often chose to mass practice of items judged as most difficult and chose to space practice of items judged as easier. By contrast, Benjamin and Bird reported that participants more often chose to mass practice of easy items and to space practice of difficult items. Our primary goals in the present article are to resolve the inconsistency between these two studies and

to evaluate theoretical hypotheses related to how students make allocation decisions when learning information. To foreshadow, across three experiments, we replicated the pattern of results both from Son and from Benjamin and Bird. Concerning our theory-relevant goal, we tested a number of hypotheses that were first proposed by Son: the massing hypothesis, the spacing hypothesis, and the metacognitive hypothesis. We describe each below and provide supporting data when possible. Finally, we discuss potential explanations for why the results of the two previous studies were inconsistent and how we explored them in the present experiments.

Theory-Based Hypotheses for How People Decide to Mass Versus Space Practice

The massing hypothesis states that people will choose to mass practice because they believe that memory is better after massed than after spaced practice (Son, 2004). Evidence suggesting that students believe that massed practice is superior (or at least as effective as spaced practice) has come from research about people's judgments of learning (JOLs), which are predictions about the likelihood that one will be able to remember a given item on a later test. For instance, Simon and Bjork (2001) had participants make JOLs after massed versus spaced practice in a motor-learning study. JOLs were higher after massed than after spaced practice, indicating that the participants believed that information was learned better after massed than after spaced practice. Similarly, Dunlosky and Nelson (1994) had college students make immediate JOLs

for noun–noun word pairs after either massed or spaced practice. No differences occurred in the magnitude of immediate JOLs between the massed and spaced groups, indicating that the participants were unaware of the benefits of spaced practice (see also Kornell & Bjork, 2008; Pashler, Zarow, & Triplett, 2003). In such cases, people may choose to mass practice because they believe that it is as good as or better than spaced practice (e.g., Baddeley & Longman, 1978).

The spacing hypothesis states that people prefer to space practice of to-be-learned items. People may choose to space practice for at least two reasons. First, because memory is better after spaced than after massed practice, people may realize that an optimal strategy would be to space practice (Son, 2004). This hypothesis indicates that people's metamemory—in this case, beliefs about their memory—drives allocation decisions. Although it is possible, this particular claim about people's memory beliefs is not supported by the aforementioned results, which suggest that people believe that massed practice is at least as good as spaced practice. In contrast to the spacing hypothesis, which focuses on metamemorial beliefs, people may decide to space practice as a default, which we refer to as the *spacing-as-default* hypothesis. More specifically, people may not think that it is necessary to study an item immediately (i.e., mass practice) because they had just studied it. Thus, they choose to space practice because it is the only other option available. Note that this hypothesis does not imply that the participants are aware of the memorial benefits of spacing practice; instead, the participants choose to space practice because they do not want to continue studying a particular item. For future discussion of these hypotheses, the spacing hypothesis will refer to metacognitively driven spacing decisions, as was proposed by Son, whereas the spacing-as-default hypothesis will refer to nonmetacognitively driven spacing decisions.

The metacognitive hypothesis states that people allocate practice differentially on the basis of how well they believe an item has been learned after the initial study trial. More specifically, as was proposed by Son (2004), it seems likely “that people may control their spacing decisions differentially from item to item on the basis of their metacognitive knowledge of each item (e.g., how difficult the item is to learn)” (p. 601). So, if a participant evaluates memory for an item and decides that it has been well learned, “waiting a while before restudying that item might be an obvious strategy” (Son, 2004, p. 601). By contrast, if a participant believes that an item has not been learned well, continuing to study in a massed fashion may be beneficial. Given that, by definition, learning will be slower for difficult items than for easier ones, a prediction is that people will more often choose to space practice of easier items and mass practice of the more difficult ones.

Status of the Hypotheses and the Present Approach

The results reported by Son (2004) support the metacognitive hypothesis. Participants made allocation decisions after an initial study trial with items; they were in-

structed to choose to mass practice, space practice, or drop items from the list of words to be restudied. In addition, the participants made JOLs for each item during practice. Most relevant now, the participants more often chose to mass practice of items judged as less well learned than of those judged as more well learned, and they chose to space practice of items judged as more well learned. This evidence led Son to prefer the metacognitive hypothesis over the massing and spacing hypotheses. Benjamin and Bird's (2006) findings appear inconsistent with Son's results and with the metacognitive hypothesis. In Benjamin and Bird's experiments, initial study with each item was followed by an allocation decision in which participants were required to allocate spaced practice to half of the items and massed practice to the other half of the items. The participants more often spaced practice of difficult items and massed practice of easy ones.

Although their results appear to be in conflict, many differences do exist between the two investigations. As is shown in Table 1, their methods differed in allocation constraints, the initial presentation rate of items, the use of dropout during practice, and the materials used. Thus, we systematically manipulated each of these variables across three experiments, with our major goals being to tease apart the discrepancy between the previous studies and to use the present data to constrain hypotheses regarding how students make decisions to mass or space practice. Our analytic strategy—with respect to inferential analyses and data presentation—was modeled after the two benchmark studies in this area, so as to make comparisons between them easier. Moreover, given that our research is hypothesis driven, we chose to conduct a priori planned comparisons where appropriate (Judd & McClelland, 1989). In particular, for replications of Son (2004) and Benjamin and Bird (2006), we conducted a priori planned comparisons ($p < .05$, one-tailed). For all other comparisons, we used the Bonferroni correction to correct for possible Type I error rate from conducting multiple t tests (e.g., Cohen, 1990). Not only did we find a general preference to space practice across all three experiments, but we also replicated the results both from Son and from Benjamin and Bird. The results from each experiment are discussed in relation to the aforementioned hypotheses.

EXPERIMENT 1

In Experiment 1, we manipulated allocation constraints to evaluate whether they were responsible for the aforementioned discrepancy. Participants were assigned either

Table 1
Methodological Differences Between Son (2004)
and Benjamin and Bird (2006)

Method	Son (2004)	Benjamin and Bird (2006)
Allocation constraints	free response	split decision
Initial presentation rate	1 sec	5 sec
Dropout	yes	no
Materials	synonym word pairs	noun–noun word pairs

to a split-decision group (in which half of the items were required to be allocated to massed practice and the other half were required to be allocated to spaced practice) or to a free-response group (in which no constraint was placed on how many items could be allocated to massed or spaced practice). We used Benjamin and Bird's (2006) materials and did not include the dropout option as an allocation choice. If allocation constraints are responsible for the discrepancy, we should replicate the results of Benjamin and Bird, as well as those of Son (2004). Specifically, when the participants are required to make a split decision (i.e., one half of the items massed and the other half of the items spaced), they should choose to space practice of difficult items (as in Benjamin & Bird, 2006). When the participants can freely choose how to allocate study, they should choose to mass practice of difficult items (as in Son, 2004).

Method

Participants and Design. A total of 68 undergraduate students enrolled in an introductory psychology course at Kent State University participated in return for course credit. The participants were randomly assigned to one of two groups, split decision or free response, with 34 participants in each group.

Materials. The materials included 48 word pairs from Benjamin and Bird (2006). Each of the 48 cue words (e.g., *cake*) had both an easy target (e.g., *frosting*) and a difficult target (e.g., *blanket*) associated with it. Twenty-four of the cues were randomly assigned to be presented with their easy targets (e.g., *cake-frosting*), and the other 24 cues were assigned to be presented with their difficult targets (e.g., *cake-blanket*) (as in Benjamin & Bird, 2006).

Procedure. All task instructions and items were presented via computer. The order of presentation of the items was randomly assigned for each participant. Each item received an initial 5-sec presentation. After this initial presentation, the participants were instructed to make an allocation decision for their next study trial with that item. They were asked to decide whether to "study now" (mass practice) or to "study later" (space practice) for each item immediately after the initial study trial with that item. The participants in the split-decision group were constrained to allocate "study now" to 24 items and "study later" to the other 24 items (as in Benjamin & Bird, 2006). A counter on the screen indicated the number of items allocated to "study now" and to "study later," which allowed the participants to easily monitor the number of choices that they had remaining for each allocation type. The participants in the free-response group could allocate as many of the 48 items to "study now" or "study later" as they wished (as in Son,

2004). If they chose "study now," the word pair was immediately shown again for a 3-sec study trial. If they chose "study later," the word pair was shown for a 3-sec study trial after initial study with all other items. After all items allocated to spaced practice had been restudied, the participants completed a cued-recall final test for all items.

Results and Discussion

Final recall performance was not relevant to evaluating the a priori predictions, and any spacing effects on recall performance should be interpreted with caution, given that individual participants' decisions to space or mass practice may yield item effects. Nevertheless, for completeness and to establish that recall performance is greater after spaced than after massed practice (as was demonstrated by previous research), we present the mean proportion of items correctly recalled at final test in Table 2. The spacing effect was manifest in all conditions, and final recall performance was greater for normatively easy than for difficult items.

Most important, consider the participants' allocation decisions. For the split-decision group (Figure 1, left panel), the participants more often chose to space practice of difficult items [$t(33) = 1.92$], replicating the results of Benjamin and Bird (2006). For the free-response group (Figure 1, right panel), no differences emerged in allocation decision as a function of item difficulty ($t_s < 0.19$). However, the participants chose more often to space practice than to mass practice for difficult items [$t(33) = 2.78$], and they showed a similar pattern for easy items, although this effect failed to reach statistical significance with the Bonferroni correction [$t(33) = 2.13, p = .04$]. Thus, we failed to replicate Son's (2004) results, in which participants more often chose to mass practice of difficult items. These results are inconsistent with the metacognitive hypothesis, which states that participants will mass practice of difficult items and will space practice of easier items. The participants chose to space practice for a vast majority of items when they were allowed to allocate items to massed and spaced practice at their discretion. Because we failed to replicate Son's results, allocation constraints (the ability to allocate choices freely or to allocate them constrained to half massed practice) cannot be entirely responsible for the discrepancy between the previous studies.

Table 2
Mean Proportions of Items Correctly Recalled on Final Test, Experiments 1, 2, and 3

Group	1 sec								5 sec							
	Easy				Difficult				Easy				Difficult			
	Mass		Space		Mass		Space		Mass		Space		Mass		Space	
	M	SEM	M	SEM	M	SEM	M	SEM	M	SEM	M	SEM	M	SEM	M	SEM
Experiment 1																
Split decision									.53	.05	.73	.04	.26	.05	.49	.05
Free response									.57	.06	.81	.03	.21	.04	.50	.04
Experiment 2																
No dropout	.41	.07	.64	.04	.12	.04	.38	.05	.59	.08	.74	.04	.34	.07	.51	.05
Dropout	.46	.08	.72	.05	.17	.05	.46	.05	.54	.09	.80	.05	.23	.07	.49	.05
Experiment 3																
Dropout	.14	.03	.20	.05	.03	.01	.10	.04	.10	.04	.32	.03	.03	.01	.15	.03

Note—The 1-sec presentation rate was not used in Experiment 1.

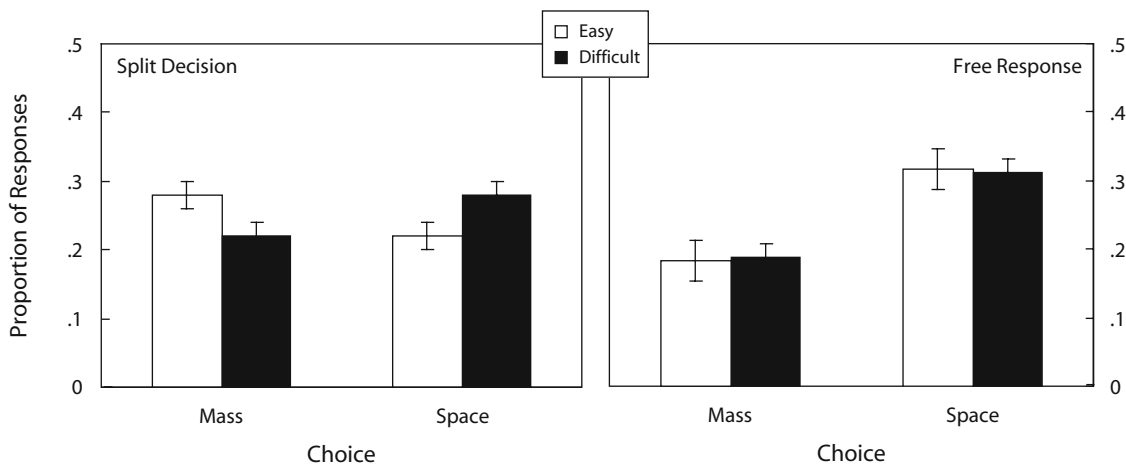


Figure 1. Mean proportions of items allocated to massed and spaced practice as a function of item difficulty (Experiment 1). The split-decision condition required half of the items to be allocated to spaced practice and half of the items to be allocated to massed practice. The free-response condition allowed items to be allocated to massed and spaced practice at the discretion of the participant. Error bars represent 1 standard error of the mean.

EXPERIMENT 2

Although we replicated the results from Benjamin and Bird (2006), we failed to replicate those from Son (2004). However, we also did not use all aspects of her design; therefore, more parametric work was needed to discover the factors responsible for the discrepancy between these two studies. The primary goal of Experiment 2 was to more closely replicate Son's design. Our choice of initial presentation rates may have been responsible for our inability to replicate Son's results in Experiment 1. In Experiment 1, the initial presentation rate was 5 sec (as in Benjamin & Bird, 2006), whereas the rate was only 1 sec in Son's experiments (see Table 1). In accord with the metacognitive hypothesis, this difference may critically influence allocation decisions. For the 5-sec rate used in Experiment 1, the participants may have preferred spacing both difficult and easy items because they believed that further immediate study would not benefit learning. By contrast, with a short initial presentation rate, participants may choose to mass practice for difficult items because these items have not been well learned, and the participants believe that they will benefit from more massed practice (cf. Metcalfe & Kornell, 2003). In Experiment 2, we evaluated these predictions by manipulating the initial presentation rate (1 vs. 5 sec).

One other aspect of the design of Experiment 2 is notable. Son (2004) included a dropout option in which participants could drop items from practice. Thus, we also manipulated dropout option: Some participants could drop items from study, whereas other participants did not have this option. If the discrepancy described above is due to a combination of both initial presentation rate and the availability of a dropout option, the short-presentation group with the dropout option should replicate the results of Son, which would represent the closest replication of her design in our experiments to this point.

Method

Participants, Design, and Materials. A total of 131 undergraduate students enrolled in an introductory psychology course at Kent State University participated in return for course credit. The participants were randomly assigned to one of four groups: short presentation dropout ($n = 33$), short presentation no dropout ($n = 33$), long presentation dropout ($n = 32$), and long presentation no dropout ($n = 33$). The materials were the same as those used in Experiment 1.

Procedure. The procedure was identical to that of Experiment 1, except for the following differences. First, all of the participants could freely choose how to allocate their practice. The participants in Benjamin and Bird (2006) had to split their decisions evenly across massed and spaced practice, which was critical for their specific goals of evaluating two current hypotheses of self-paced study. However, our main goal was to evaluate how students freely allocate their practice, which better represents typical study conditions (Dunlosky, Bottiroli, & Hartwig, 2009). Thus, a split-decision group was not included in Experiment 2. Second, the initial presentation rate of items was either short (1 sec) or long (5 sec), depending on the group to which the participants were assigned. Third, the participants assigned to the drop option groups had an additional allocation choice of dropping practice with an item. If the participants in this group chose to drop an item from study, it was not presented again. As in Experiment 1, the restudy trial (either massed or spaced practice) for each item was 3 sec.

Results and Discussion

We first present the data for groups that did not have a dropout option during practice to compare the results from Experiment 2 with those from Experiment 1. We then present the data for groups that did have the dropout option during practice. For completeness, we report the mean proportion of items correctly recalled at final test in Table 2. As in Experiment 1, the spacing effect arose in all conditions, and recall performance was greater for normatively easy than for difficult items.

No dropout option. For the short-presentation group (Figure 2, left panel), although relatively few items were allocated to massed practice, the participants chose to

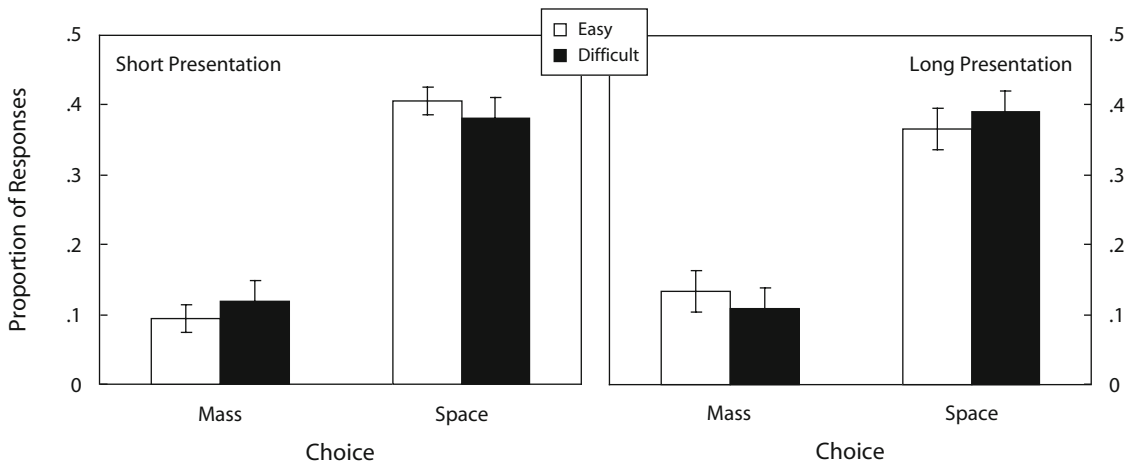


Figure 2. Mean proportions of items allocated to massed and spaced practice for no-dropout-option groups as a function of item difficulty (Experiment 2). *Short and long presentation* refer to the initial presentation rates of 1 and 5 sec, respectively. Error bars represent 1 standard error of the mean.

mass practice more often for difficult than for easy items [$t(32) = 2.91$], which is in line with the results reported by Son (2004). Importantly, the participants most often chose to space practice for both easy and difficult items ($ts > 5.00$).

For the long-presentation group (Figure 2, right panel), no differences emerged in choice of allocation (mass vs. space) as a function of item difficulty. However, for both easy and difficult items, the participants overwhelmingly chose spaced practice over massed practice ($ts > 3.81$). Taken together, these results suggest an overall preference to space practice for both easy and difficult items.

Dropout option. The short-presentation group with the dropout option (Figure 3, left panel) is a methodological replication of that of Son (2004), with the exception that we used Benjamin and Bird's (2006) materials. We expected that allocation decisions would differ on the basis of item difficulty. Most important, as is shown in Figure 3

(left panel), the participants chose to mass practice more for difficult than for easy items [$t(32) = 2.97$], consistent with the results reported by Son. Of those items allocated to be dropped from practice, the participants chose to drop easier items more often than difficult items [$t(32) = 3.34$]. Finally, for both easy and difficult items, a trend showed participants more often chose to space than to mass practice ($ts > 2.50$, $ps < .018$).

The long-presentation group (Figure 3, right panel) methodologically differs from the previous group only by the initial presentation rate. Thus, if presentation rate is not an important variable, we should replicate the results of the short-presentation group that had the dropout option. Of the few items allocated to massed practice, no differences emerged as a function of item difficulty [$t(31) = 0.38$]. Of the items allocated for spaced practice, the participants more often chose to space practice for difficult items than for easy items [$t(31) = 5.62$]. Of the items

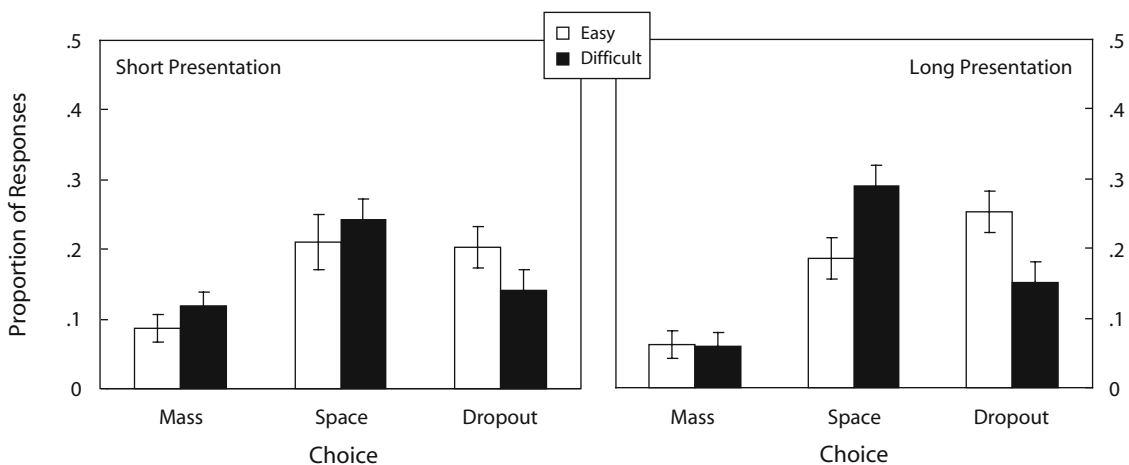


Figure 3. Mean proportions of items allocated to massed, spaced, and dropout practice for dropout-option groups as a function of item difficulty (Experiment 2). *Short and long presentation* refer to the initial presentation rates of 1 and 5 sec, respectively. Error bars represent 1 standard error of the mean.

allocated to be dropped from practice, the participants more often chose to drop easier items than difficult items [$t(31) = 5.92$]. Finally, for both easy and difficult items, the participants more often chose to space practice than to mass practice ($ts > 3.28$). In summary, these results indicate an overall preference for spacing practice for both easy and difficult items, which is consistent with the results from Experiment 1.

EXPERIMENT 3

The results from Experiment 2 suggest that initial presentation rate is partially responsible for the discrepancy between the results in Experiment 1 (as well as those of Benjamin & Bird, 2006) and those in Son (2004). Nevertheless, as compared with the participants in Son, those in Experiment 2 were much less likely to mass practice of difficult items overall: In Son, the participants massed 51% of the difficult items, whereas in Experiment 2 here, only 24% of these items were massed. Thus, in Experiment 3, we wanted to consider two other design differences across studies that could have resulted in this discrepancy and that could potentially moderate the degree to which participants choose to mass (or space) difficult items.

First, one major difference in the design of Experiment 2 and the one used by Son (2004) was the kind of material used (Table 1), which does provide a viable explanation for the present discrepancy. In particular, the items used in Experiments 1 and 2 (from Benjamin & Bird, 2006) and those used by Son appear to differ in difficulty. For example, *jungle–chimney* was a paired associate used by Benjamin and Bird, whereas *obstreperous–unruly* was a pair from Son. The items in the former study comprised familiar words that all the participants would know the meanings of prior to the experiment, whereas those in the latter included abstract words that at least some of the participants would likely not know. In terms of the metacognitive hypothesis, the synonym word pairs may be difficult for the participants to remember after the initial study trial, and hence, they may believe that practicing

them immediately would benefit learning. Thus, in Experiment 3, we used Son's materials to evaluate whether this variable influences allocation decisions.

Second, Son (2004) also had her participants make a JOL immediately after they studied each item (e.g., *obstreperous–unruly*) during the initial presentation. JOLs are people's predictions about the likelihood of being able to later recall the response (i.e., *unruly*) when shown the corresponding stimulus (i.e., *obstreperous*) on the criterion test. Son used the JOLs to scale item difficulty, and hence, to best match her design and analytic method, we also had our participants make JOLs in Experiment 3. Moreover, we had an independent group of students rate Son's items on learning difficulty, and we used these ratings to choose subsets of the most difficult and easiest items, so that we could also conduct analyses as in our previous experiments and as in Benjamin and Bird (2006).

Method

Participants and Design. A total of 72 undergraduate students enrolled in an introductory psychology course at Kent State University participated in return for course credit. The participants were randomly assigned to one of two groups, short presentation or long presentation, with 36 participants in each group.

Materials and Procedure. Sixty synonym word pairs from Son (2004) were used (e.g., *brook–tolerate*). The procedure was identical to that of Experiment 2, except for the following differences. First, in Experiment 3, all of the participants had the dropout option in addition to the mass and space options. Second, before making an allocation decision, the participants were asked to make a JOL for each word pair by indicating the likelihood that they would be able to recall the word pair on a later test (but they were not specifically told when the test would occur). The participants were asked to type in any number from 0 to 100 to indicate the likelihood of later retrieval.

Results and Discussion

As in Experiments 1 and 2, the mean proportion of items correctly recalled at final test is reported in Table 2.

Short-presentation group. The short-presentation group (Figure 4, left panel) is an exact methodological replication of Son (2004), except that the participants were drawn from different participant pools. In contrast

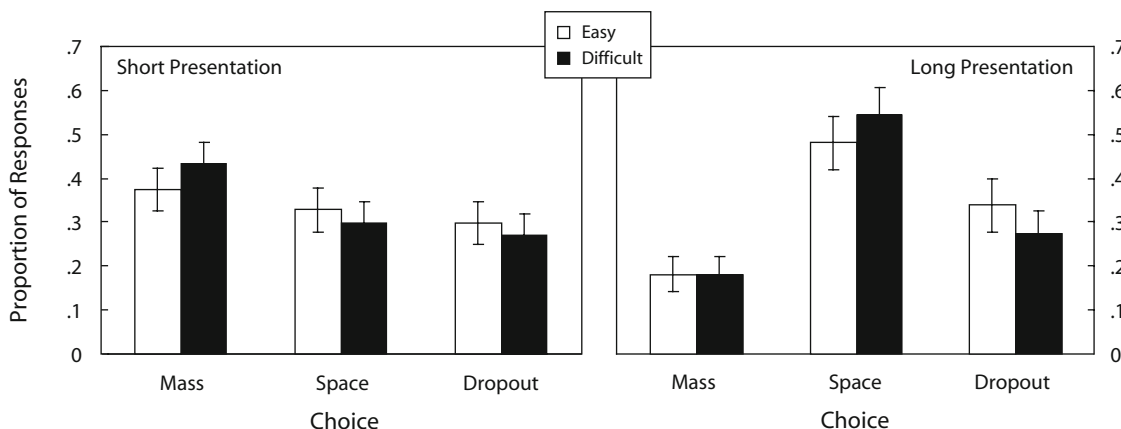


Figure 4. Mean proportions of items allocated to massed, spaced, and dropout practice as a function of item difficulty (Experiment 3). *Short and long presentation* refer to the initial presentation rates of 1 and 5 sec, respectively. Error bars represent 1 standard error of the mean.

to the results of Experiments 1 and 2, many more items were allocated to massed practice. The participants more often chose to mass practice for difficult items than for easy items [$t(35) = 2.34$], which directly replicates the results of Son. Of the items allocated for spaced practice, no reliable differences emerged for a preference of spacing practice (or dropping practice) of easy versus difficult items. In contrast to Experiments 1 and 2, with a short presentation rate (1 sec), the participants showed no general preference to space practice; instead, they tended to prefer to mass practice.

We also analyzed the data for this group on the basis of the participant JOLs as a criterion for item difficulty. JOL values were vincentized into six levels for each participant (Figure 5). JOL level is of primary interest here, since these are the items that the participants judged as most difficult. The participants overwhelmingly chose to mass practice for the most difficult items [JOL Level 1: mass vs. space, $t(35) = 3.73$; mass vs. dropout, $t(35) = 5.47$]. The difference between massed, spaced, and dropout practice for these most difficult items is quite striking and clearly replicates the results reported by Son (2004).

Long-presentation group. For the long-presentation group (Figure 4, right panel), of the few items allocated for massed practice, the participants did not choose more often to mass practice for easy than for difficult items. Of the items allocated for spaced practice, the participants more often chose to space practice for difficult items than for easy items [$t(35) = 3.65$]. Of the items allocated to be dropped from practice, the participants more often chose to drop out practice for easier items than for difficult items [$t(35) = 3.35$]. Finally, for both easy and difficult items, the participants more often chose to space practice than to mass practice ($t_s > 3.23$). Thus, the participants preferred to space rather than to mass practice for both easy and difficult items with a long initial presentation rate (replicating the results from Experiments 1 and 2), but not with a short initial presentation rate.

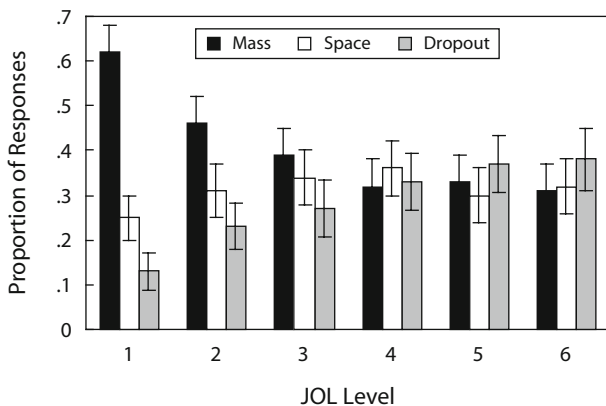


Figure 5. Mean proportions of items allocated to massed, spaced, and dropout practice as a function of subjective item difficulty (Experiment 3). Judgment-of-learning (JOL) ratings were split into six groups on the basis of JOL judgments made by the participants, from lowest JOL (1, *most difficult*) to highest JOL (6, *easiest*). Error bars represent 1 standard error of the mean.

GENERAL DISCUSSION

Across three experiments, participants showed a general preference to space practice for both easy and difficult items, although key exceptions arose. The participants more often massed practice of easy items (and spaced practice of difficult items) when they were forced to mass practice of half of all items (i.e., split-decision group, Experiment 1), which replicated the results reported by Benjamin and Bird (2006). Also, the participants preferred to mass practice when difficult materials and a short initial presentation rate (1 sec) were used (Experiment 3), which replicated the results reported by Son (2004). Recently, we were made aware of an independent investigation from Toppino, Cohen, Davis, and Moors (2009), who examined the same issue. The design of their first experiment and that of our Experiment 3 are shockingly similar and, hence, unsurprisingly yielded a similar empirical outcome: When Son's difficult materials were used, the participants preferred to mass practice of difficult items when the initial presentation rate was brief, but they preferred to space practice when the initial rate was longer. With this converging evidence in mind, however, the present experiments also included a systematic analysis of the broader array of methodological differences in the previous research (Table 1), which provided further evidence about how students make allocation decisions.

Current Status of the Theoretical Hypotheses

Given that we replicated the results of both Son (2004) and Benjamin and Bird (2006), how people decide whether to mass or to space practice becomes a critical question. Broadly speaking, some of the theoretical hypotheses posed to answer this question implicate qualitatively different mechanisms for these effects; that is, they implicate either metamemory or nonmetamemory causes for allocation decisions. For the former, allocation decisions arise from people's beliefs about memory or from their use of memory monitoring during study. For the latter, allocation decisions arise from nonmetamemorial factors, such as using the space option as a default. It is vital to contrast and competitively evaluate these classes of mechanisms, because although allocation decisions are metacognitive acts (i.e., they involve controlling one's memory; Dunlosky & Metcalfe, 2009), that does not mean that they must be explicitly driven by beliefs and metamemorial processes. Thus, to facilitate further efforts to evaluate hypotheses for allocation decisions, we will situate each hypothesis within the general classes of metamemory or nonmetamemory causes. Table 3 provides an overview of these hypotheses that includes the current status of each hypothesis based on the outcomes from the present experiments.

Metamemory and nonmetamemory hypotheses. Hypotheses implicating metamemory causes for allocation decisions include the massing, spacing, and metacognitive hypotheses (Son, 2004). To revisit, the massing and spacing hypotheses state that people believe that massed (or spaced) practice is best for memory, so they will favor massed (or spaced) practice for most items. By

Table 3
Hypotheses for Why Students Decide to Space or Mass Their Practice

Hypothesis	Current Status	Relevant Data
Metamemory Hypotheses		
Massing Spacing	Disconfirmed	Ps overwhelmingly prefer to space practice (Experiments 1–3).
	Limited contribution	Ps often prefer to space, but do not believe spacing is better; Ps do not always prefer to space practice (Experiment 3).
Metacognitive	Partial contribution	In some cases, Ps did prefer to mass difficult items (Experiment 3).
Nonmetamemory Hypotheses		
Spacing as default	Partial contribution	In some cases, Ps appeared to use the spacing option because they could not drop items from study (Experiment 2).
Degraded perception	Partial contribution	Ps appear to mass practice when they have not fully perceived items (see Table 4 and the General Discussion section).

Note—See the text for discussion of the relative support for these hypotheses. Ps, participants.

contrast, the metacognitive hypothesis states that participants' beliefs about how to best enhance memory are used to make allocation decisions (Son, 2004). When items are not well known, participants believe that massed practice is best for memory (e.g., mass practice for difficult items), whereas when items are relatively well known, they believe that spaced practice is best (e.g., space practice for easy items).

Hypotheses implicating nonmetamemory causes for allocation decisions include the spacing-as-default hypothesis and the degraded-perception hypothesis. The spacing-as-default hypothesis states that participants space practice because they are finished studying an item and opt to space practice because there is no better option. The degraded-perception hypothesis states that participants will mass practice when degraded perception at encoding causes items to not be fully processed. Variables such as distraction, lack of knowledge, presentation rate, and difficulty of reading items may disrupt processing and lead to degraded perception of the stimuli.

Hypotheses in relation to the present data. The present experiments do not fully resolve which causes (or combination of them) are responsible for people's allocation decisions, but importantly, they do offer new insight into how people make allocation decisions. For instance, as was discussed above, the results from the present experiments showed that, in general, participants more often chose to space than to mass practice. These results disconfirm the massing hypothesis, and although the spacing hypothesis provides a straightforward explanation for this preference, recall that previous research suggests that people generally believe that massed practice is at least as good for memory as is spaced practice. Thus, these two metamemory-based hypotheses do not provide a viable explanation for people's allocation decisions.

By contrast, at least part of people's preference for spaced practice arises from using this option as a default. In particular, collapsing across initial presentation rates and item difficulty (Experiment 2), the proportion of items allocated to spaced practice is lower for the group that could drop items from study than for those who could not [$t(129) = 5.38, p < .001$]. Thus, when the only options are to mass or space practice, these participants sometimes used the space option for items that they would prefer to

drop from study. Although this evidence is consistent with the spacing-as-default hypothesis (Table 3), it is possible that people sometimes use the option to mass practice as a default. For instance, in an unpublished study by Lisa Son, participants were asked why they chose to mass practice for some items. This study did not include the dropout option, and some participants reported that they chose to mass practice of easy items because they "wanted to get it over with" (Son, personal communication, 2009). Thus, in some circumstances, participants may use the mass option as a default. Our main point, then, is that people make some allocation decisions for reasons that have little—if anything—to do with metamemory causes.

As was noted above, the participants generally more often spaced than massed practice; nevertheless, they did mass practice on a substantial proportion of trials in some circumstances, which indicates that the spacing-as-default hypothesis cannot fully explain the present results. These observations led us to inquire whether decisions to mass practice are driven by metamemory or nonmetamemory causes. Consistent with a metamemory cause, the results from the short-presentation group in Experiment 3 support the metacognitive hypothesis. As is shown in Figure 5, the participants were more likely to mass pairs judged to be the most difficult (JOL Level 1) than those judged to be the easiest (JOL Level 6). Even though the metacognitive hypothesis can explain these allocation decisions (as in Son, 2004), the degraded-perception hypothesis

Table 4
Mean Proportions of Easy Items and Difficult Items Perceived After a Brief Presentation (1 sec) Based on Subjective and Objective Measures

Items	Easy		Difficult	
	<i>M</i>	<i>SEM</i>	<i>M</i>	<i>SEM</i>
Subjective Measure				
Synonym word pairs	.90	.02	.59	.02
Noun–noun word pairs	.98	.01	.99	.01
Objective Measure				
Synonym word pairs	.85	.03	.50	.06
Noun–noun word pairs	.97	.01	.96	.01

Note—Synonym word pairs (e.g., *obstreperous–unruly*) were used by Son (2004), and the noun–noun word pairs (e.g., *jungle–chimney*) were used by Benjamin and Bird (2006).

provides an alternative explanation for why participants mass practice. According to this nonmetamemory hypothesis (Table 3), participants elect to mass practice when an item is not fully perceived during the initial presentation. And in the present experiments, the participants massed practice most often when items were presumably more difficult to process during the initial presentation trial. In Experiment 3, the participants more often massed practice when items were difficult synonym pairs (e.g., *obstreperous-unruly*) that were initially presented briefly. When a longer presentation rate was used, the participants chose to mass practice much less often (Figure 4). Moreover, in Experiment 2, when items were relatively easy noun-noun pairs (e.g., *jungle-chimney*), the participants chose to mass practice relatively infrequently, even when the initial presentation rate was brief (Figure 3).

This evidence indirectly supports the degraded-perception hypothesis. To more directly evaluate it, we conducted the following miniexperiment. Thirty-one college students were shown 40 word pairs. Twenty word pairs were the synonym pairs from Son (2004), and 20 were the noun-noun pairs from Benjamin and Bird (2006). For each set of 20 word pairs, 10 were relatively easy and 10 were relatively difficult. The word pairs were presented at the short presentation rate (1 sec). After the initial presentation, the participants made a *yes-no* judgment regarding whether they were able to read both words in the pair. After making this subjective judgment, the participants then attempted to type both words on the computer, which provided an objective measure of whether they fully perceived the particular word pair. We then computed the mean proportion of times that the participants reported not being able to read the items and the proportion of times that they could not type them (subjective and objective measures, respectively; see Table 4). Consistent with the degraded-perception hypothesis, the mean proportion of items that the participants reported perceiving was lower for the synonym word pairs than for the noun-noun word pairs, and as important, the participants' perception was even further degraded for the most difficult synonym word pairs. These observations were supported by ANOVAs (conducted separately for each measure) that revealed main effects of material set (synonyms vs. noun-noun pairs) ($F_s > 55.0$, $MS_{eS} < 0.04$, $ps < .001$) and a significant interaction between item difficulty and material set ($F_s > 32.14$, $MS_{eS} = 0.03$, $ps < .001$). These data provide further evidence that the participants' decisions to mass practice with Son's materials was in part due to degraded perception of those items during the initial presentation trial.

Converging evidence for the degraded-perception hypothesis also comes from Toppino et al. (2009, Experiment 2), who also presented items at a rate of 1 sec/item; after the initial presentation of an item, participants first made an allocation decision and then reported whether they had seen both words in a pair. Toppino et al. could not examine the critical interaction involving material set (reported above), because they only used Son's (2004) materials. However, because the participants made an al-

location decision and reported whether they had seen each item, Toppino et al. could evaluate whether degraded perceptions were correlated with decisions to mass practice. As was expected, when the participants reported that they had not seen both words in a pair (vs. when they reported seeing both), they were more likely to choose to mass practice than to space practice.

Concluding Remarks

The present data indicate that nonmetamemorial causes are at least partially responsible for people's allocation decisions. The present experiments were not designed to estimate the relative contribution of all of these factors to people's allocation decisions. Nevertheless, it seems likely that the various causes (Table 3) will differentially contribute to allocation decisions depending on individual differences in students and the prevailing context of study: Individuals who have learned about the power of spaced practice may use it exclusively (as in the spacing hypothesis), the easiest (and already known) items may be spaced when no dropout option is available (as in the spacing-as-default hypothesis), and many items may be massed when the environment undermines initial processing of items (as in the degraded-perception hypothesis). We await future research to systematically explore these possibilities. Most important, however, the present results identified the factors that led to the apparent discrepancy between Son (2004) and Benjamin and Bird (2006), and we also found a general preference to space practice for both easy and difficult items. Thus, even if students are not aware of the memorial benefits of spaced practice, they do choose this relatively effective schedule of learning more often than massed practice.

AUTHOR NOTE

This project was supported by a James S. McDonnell Foundation 21st Century Science Initiative in Bridging Brain, Mind and Behavior Collaborative Award. Thanks to Cristin Clewell, Melissa Roderick, Sara Smith, and Katie Wissman for assistance with data collection. Special thanks also to Katherine Rawson and RADlab for helpful comments and feedback on this project. Correspondence concerning this article should be sent to J. Dunlosky, Department of Psychology, Kent State University, P.O. Box 5190, Kent, OH 44242-0001 (e-mail: jdunlosk@kent.edu).

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. doi:10.1080/0014013 7808931764
- BENJAMIN, A. S., & BIRD, R. D. (2006). Metacognitive control of the spacing of study repetitions. *Journal of Memory & Language*, **55**, 126-137. doi:10.1016/j.jml.2006.02.003
- 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. doi:10.1037/0033-2909.132.3.354
- COHEN, J. (1990). Things I have learned (so far). *American Psychologist*, **45**, 1304-1312. doi:10.1037/0003-066X.45.12.1304
- DEMPSTER, F. N. (1988). The spacing effect: A case study in the failure to apply the results of psychological research. *American Psychologist*, **43**, 627-634. doi:10.1037/0003-066X.43.8.627
- DONOVAN, J. J., & RADOSEVICH, D. J. (1999). A meta-analytic review of the distribution of practice effect: Now you see it, now you don't.

- Journal of Applied Psychology*, **84**, 795-805. doi:10.1037/0021-9010.84.5.795
- DUNLOSKY, J., BOTTIROLI, S., & HARTWIG, M. (2009). Sins committed in the name of ecological validity: A call for representative design in education research. In D. Hacker, J. Dunlosky, & A. Graesser (Eds.), *Handbook of metacognition in education* (pp. 430-440). New York: Psychology Press.
- DUNLOSKY, J., & METCALFE, J. (2009). *Metacognition*. Thousand Oaks, CA: Sage.
- 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 & Language*, **33**, 545-565. doi:10.1006/jmla.1994.1026
- EBBINGHAUS, H. (1913). *Memory: A contribution to experimental psychology* (H. A. Ruger & C. E. Bussenius, Trans.). New York: Teachers College, Columbia University. (Original work published 1885)
- 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). Potomac, MD: Erlbaum.
- JUDD, C. M., & McCLELLAND, G. H. (1989). *Data analysis: A model-comparison approach*. Orlando, FL: Harcourt Brace.
- KORNELL, N., & BJORK, R. A. (2008). Learning concepts and categories: Is spacing the "enemy of induction"? *Psychological Science*, **19**, 585-592. doi:10.1111/j.1467-9280.2008.02127.x
- MELTON, A. W. (1970). The situation with respect to the spacing of repetitions and memory. *Journal of Verbal Learning & Verbal Behavior*, **9**, 596-606. doi:10.1016/S00225371(70)80107-4
- 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. doi:10.1037/0096-3445.132.4.530
- PASHLER, H., ZAROW, G., & TRIPLETT, B. (2003). Is temporal spacing of tests helpful even when it inflates error rates? *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **29**, 1051-1057. doi:10.1037/0278-7393.29.6.1051
- SIMON, D. A., & BJORK, R. A. (2001). Metacognition in motor learning. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **27**, 907-912. doi:10.1037/0278-7393.27.4.907
- SON, L. K. (2004). Spacing one's study: Evidence for a metacognitive control strategy. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **30**, 601-604. doi:10.1037/0278-7393.30.3.601
- TOPPINO, T. C., COHEN, M. S., DAVIS, M. L., & MOORS, A. C. (2009). Metacognitive control over the distribution of practice: When is spacing preferred? *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **5**, 1352-1358. doi:10.1037/a0016371

(Manuscript received May 8, 2009;
revision accepted for publication November 12, 2009.)