Available online at www.sciencedirect.com
SCIENCE DIRECT®
Journal of Memory and Language 55 (2006) 126–137
Journal of Memory and Language
www.elsevier.com/locate/jml

Metacognitive control of the spacing of study repetitions ☆

Aaron S. Benjamin *, Randy D. Bird

Department of Psychology, University of Illinois at Urbana-Champaign, 603 East Daniel Street, Champaign, IL 61820, USA

Received 25 November 2005; revision received 11 February 2006
Available online 18 April 2006

Abstract

Rememberers play an active role in learning, not only by committing material more or less faithfully to memory, but also by selecting judicious study strategies (or not). In three experiments, subjects chose whether to mass or space the second presentation of to-be-learned paired-associate terms that were either normatively difficult or easy to remember, under the constraint that subjects needed to space exactly half of the items (and mass the other half). In contrast with recent findings that implemented no such constraint (Son, 2004), subjects chose to space more of the difficult pairs (in Experiments 1 and 2). Reduction in exposure time eliminated but did not reverse this effect (Experiment 3). Subjects who spaced more of the difficult pairs were more likely to exhibit superior memory performance, but, because subjects who made spacing selections that had no effect on the actual scheduling of items also showed this effect (Experiment 2), that enhancement in performance is more likely to reflect subject selection than strategy efficacy. Overall, these results suggest that choice constraints strongly elicit a discrepancy-reduction approach (Dunlosky & Hertzog, 1998) to strategic decision-making, but that reduced study time can eliminate this effect.

Keywords: Memory; Metamemory; Judgments of learning; Spacing effect; Self-regulated study

The question of how people use their metacognitive knowledge to regulate their behaviors has been of much interest in recent years, particularly with regard to the implementation of study strategies. Metacognition plays an integral role in tasks such as self-directed learning (Koriat & Goldsmith, 1996; Nelson & Narens, 1990), and understanding the means by which metacognitions guide learning processes is essential to facilitate and optimize the learning process itself (see Bjork, 1994).

This article pursues that question by examining the strategies that subjects employ in the scheduling of learning events. Recent evidence revealed conditions under which subjects prefer to space easier materials and mass more difficult ones (Son, 2004). That result is fascinating because it either reveals that subjects choose to apply more effective study conditions to easier materials—a result in conflict with the vast majority of findings from study-time allocation experiments—or it reveals a fundamental misappreciation of the greater effectiveness of spacing in promoting learning (e.g., Baddeley & Longman, 1978). However, the present experiments reveal the opposite effect—subjects choose to space difficult and mass easy items. These results thus

☆ The authors wish to thank Lisa Son and Juhn Dunlosky, as well as members of the Human Memory and Cognition Laboratory at the University of Illinois, for helpful discussions and commentaries on this article. Portions of this work were presented at the 45th Annual Meeting of the Psychonomic Society. Correspondence should be addressed to Aaron Benjamin at ASBENJAM@UIUC.EDU.
* Corresponding author. Fax: +1 217 244 5876.
E-mail address: ASBENJAM@CYRUS.PSYCH.UIUC.EDU (A.S. Benjamin).

0749-596X/$ - see front matter
doi:10.1016/j.jml.2006.02.003

suggest that, under some conditions, subjects do understand the beneficial effects of spacing and also choose to selectively utilize them with difficult materials.

Self-regulation of learning

Theories of self-regulated study claim that active learners use assessments of item difficulty and their own degree of learning in deciding whether to allocate further cognitive resources toward study of that item or to move on to other items (e.g., Dunlosky & Hertzog, 1998; Mazzoni, Cornoldi, & Marchitelli, 1990; Metcalfe, 2002; Nelson & Leonesio, 1988; Nelson & Narens, 1990; Thiede & Dunlosky, 1999). There is some debate, however, with regard to *how* difficulty and resource allocation—specifically, study time allotment—are related.

Discrepancy reduction

One theory emphasizes a discrepancy-reduction mechanism (Dunlosky & Hertzog, 1998). According to this theory, the learner compares their perceived degree of learning of a to-be-learned item to the their desired level of mastery for that item, also known as the norm of study (Le Ny, Denhiere, & Le Taillanter, 1972; see also Nelson & Narens, 1990), and if the degree of learning does not reach that criterion, additional study of that item ensues. Therefore, in reducing the discrepancy between an item's current and desired degree of learning, the model predicts an inverse relationship between the perceived (prior) degree of learning and study time, and hence suggests that people will allot more study time to judged-difficult than judged-easy items (Dunlosky & Hertzog, 1998; see also Thiede & Dunlosky, 1999).

Indeed, a multitude of experiments have shown that people tend to study more difficult items for longer than they study easier items. In a comprehensive review, Son and Metcalfe (2000) reported that, of 46 treatment combinations in 19 published experiments in which subjects controlled the allocation of study time, 35 revealed a strategy of devoting more time to difficult items, and none showed the opposite strategy of devoting more time to easy items. These studies included subjective as well as objective measures of difficulty, and the results were consistently found across age groups and study materials.

Proximal learning

However, Son and Metcalfe (2000) showed that total time constraints caused subjects to apportion more study time to judged-easy items than to judged-difficult items. That is, when the total study time allotted was likely insufficient to master all items, subjects chose to allocate their limited time to items that individually take less time to master, rather than slowly learn fewer difficult items. Similarly, Thiede and Dunlosky (1999) found that if their task was to remember only a small portion of the to-be-learned items, rather than the full set, subjects devoted more study time to easy items.

Metcalfe (2002) surmised that the discrepancy-reduction model adequately accounted for subjects' allocation strategies only under certain conditions, and forwarded a more comprehensive model to incorporate the newer data. She argued that study time should be devoted to those items that are just beyond the current grasp of the individual, in the region of *proximal learning* (Metcalfe, 2002; see also Metcalfe & Kornell, 2003). In the case that those just-unlearned items are the most difficult to-be-learned items, the discrepancy-reduction and region of proximal learning models agree on what the appropriate strategy should be. However, in cases where easy items are still unlearned, the predictions of the two theories are in opposition.

Whereas the discrepancy-reduction model suggests that learners will always devote more time to the difficult items, the proximal learning hypothesis implies that individual differences in expertise within a domain should influence study-time allocation. Metcalfe (2002) demonstrated this effect using English-Spanish vocabulary pairs. To monolingual speakers, even relatively easy items can be difficult to learn, and thus those speakers allocated more study time to those easy pairs accordingly. Experts, on the other hand, spent more time studying the difficult word pairs, and Metcalfe (2002) attributed the group differences in item selection to the difference between the two groups' regions of proximal learning. Novices chose to spend more time studying the easy, yet still unlearned, items before moving on to more difficult items, a result that is not predicted by the discrepancy-reduction model.

Effects of strategy choice

In addition to identifying the strategies used in allocating study time, determining which strategy ultimately leads to superior performance on subsequent recall tests is of importance as well. The actual quality of any study strategy, after all, can only be evaluated by the outcome it produces on the subsequent test. As Son and Metcalfe (2000) pointed out, even though much of the previous literature suggests a tendency for subjects to study the difficult items for longer than the easy items, there are no data showing that subjects who employ such a strategy outperform subjects who spend equal amounts of time on easy and difficult items. While it is intuitive that increased duration of study should lead to higher recall performance for any given item, previous findings have suggested that once study time for an item has been sufficient for initial acquisition, continued immediate study of that item leads to little or no increase in the probability of future recall. This null increase in performance despite substantial increases in study time has been termed the "labor-in-vain effect" (Nelson & Leonesio, 1988, p. 680).

Metcalfe and Kornell (2003) systematically investigated the effects of allocating study time to a particular subset of items on later recall performance. Their results showed that allocating study time to items of medium difficulty was more helpful than was allocating study time to easy or difficult items, presumably because learning of the easy items had already plateaued and would not benefit from additional study, whereas learning of the difficult items would require even greater amounts of additional study time to reach a level sufficient to improve later recall. The plateau occurs sooner for easy items than for medium items, and likewise sooner for medium items than for difficult items.

If an optimal strategy of study-time allocation is one in which study of a particular item should be discontinued at the point when further study is not as advantageous as study of another item, a critical question for the metacognizer is where (or when) that point is. If information uptake has slowed sufficiently for an item, immediate further study of that item at the cost of a different item may be detrimental to overall recall performance, whereas further study at some later time may boost one's memory for that item and may benefit future recall. Such a strategy has already been shown in the context of semantic retrieval tasks in which subjects tried to name as many items as possible from two semantic categories in a short time. When retrieval for items in one category slowed, and the other category became relatively more appealing, subjects switched to the second category (Young, 2004). To test whether subjects employ this type of strategy in a learning task, a paradigm can be used that holds total study time constant by taking advantage of the well-known effects of spacing (e.g., Crowder, 1976).

Spacing and strategy selection

In paradigms in which to-be-learned items are presented multiple times, recall performance is typically bolstered when those presentations are temporally spaced apart, as opposed to massed together. This *spacing effect* has been found in a variety of learning tasks using many different materials (for reviews, see Crowder, 1976; Dempster, 1996). Until recently, there had been very few reported exceptions to this rule. Glenberg (1977) found that massing is advantageous over spacing if the recall test immediately follows study termination, but this result likely reflects the proximity of both study presentations to the test overshadowing the advantages of spacing.

Metcalfe and Kornell (2003) recently illustrated another limitation of the spacing effect by using very short presentation times during study. They found that, with items of a medium difficulty level, the advantage of spacing over massing disappeared at presentations of 1 s, and even became disadvantageous at presentations of 0.5 s. Presumably, after such a short study time, learning of the item had not plateaued such that immediate further study would be "labor in vain." Thus, they posited that the spacing effect only obtains when to-be-studied items are sufficiently processed before termination of the first presentation.

Unlike study-time allocation studies, in which subjects often hold control of some aspect of the study list, experimenters have almost always controlled study lists in examinations of the spacing effect, with one recent exception. Son (2004) hypothesized that, when given the choice of massing or spacing each item, subjects would implement a metacognitive control strategy to determine which re-study option would be most beneficial for each item. While the literature indicates that spacing all items would be an actual optimal strategy, it also shows that subjects commonly overestimate their own degree of learning for study materials studied under massed conditions (e.g., Shaughnessy, 1976; Zechmeister & Shaughnessy, 1980). Similarly, Baddeley and Longman (1978) demonstrated that many subjects preferred massed over distributed practice. Thus, massing all items may seem to be an optimal strategy from a subject's perspective, despite the objective evidence to the contrary. Barring these two extreme strategies, Son (2004) hypothesized that subjects' decisions of which items to mass and which items to space would be guided by their metacognitive judgments of how difficult each item would be to recall, as measured by judgments-of-learning (JOLs; for a review, see Schwartz, 1994).

In Son's (2004) experiment, subjects were allowed to mass, space, or terminate their study of each item, with no constraints on the assignment of those three options. In general, subjects chose to space items more often than to mass items. Subjects also massed more of the judged-difficult items than judged-easy items. That is, they appeared to employ a strategy of immediately restudying items that were perceived as difficult or unlearned, while delaying re-study or terminating study of items perceived as easy or sufficiently learned. Such a strategy is, on the face of it, inconsistent with the majority of results from the study-time allocation studies, because the more difficult items were not allocated to the more effective study condition. From a discrepancy-reduction standpoint, this finding is thus likely suboptimal (although strong claims about the optimality of one strategy or another are highly dependent on the initial state of knowledge, as discussed earlier). Interestingly, despite their clear preference to mass judged-difficult items, subjects in Son's (2004) experiment recalled a higher percentage of spaced items than massed items at all levels of item difficulty (that is, across the JOL spectrum).

Because Son's (2004) findings are inconsistent with a discrepancy-reduction hypothesis, and thus have major implications for the viability of that hypothesis in self-scheduling paradigms, the goal of the present experiments was to evaluate whether those findings generalize to conditions that might encourage subjects to employ a discrepancy-reduction strategy and that also decrease the effects of experimental factors that work against the use of that strategy. Thus, there are several major methodological differences between her experiment and ours.

The first set of changes concerns the selection procedure and consequent sources of variance. Subjects in Son's (2004) could have massed all items or spaced all items, had they so desired. We chose to enforce limits on the proportion of items an individual subject could mass and space for several reasons. First, it eliminates individual differences in the proportion of items chosen for a particular regimen. This is important because it allows us to assert that the basic effects we are interested in—namely, *individual* choices of study regimen for easy and difficult items—are not conflated with group or subgroup effects. For example, it might be the case that the smartest subset of subjects would indeed choose to space difficult items and mass easy ones, but are also smart enough to space all of the items when given that opportunity. The second advantage of controlling the relative proportions of spacing selections is that it may force subjects to use their desired choices more sparingly. If subjects consider the option of spacing a limited resource, it stands to reason that they will be more deliberate about when they choose to implement that option. Just as total time constraints cause subjects to allocate study time differently than if are provided unlimited time (Metcalfe, 2002; Son & Metcalfe, 2000), it seems likely that subjects given limited massing and spacing opportunities may also allocate them differently than if they had no such restrictions. Finally, limiting opportunities for spacing further serves the goal of approximating real-world study constraints in which limited time prior to a test necessarily implies tradeoffs in what materials one studies at a particular time.

A second major change involves the elimination of the ''Done'' option for restudy. In Son's (2004) experiment, subjects had the option of not restudying the item at all. Having that option introduces two related sources of variance that we wished to exclude. First, because different numbers of items will drop out of the analysis for different subjects, subjects will contribute differentially to group means. Second, certain types of items will be more likely to drop out than others, thus introducing an item confound. These two effects also combine to yield an interactive subject $\times$ item confound as well.

However, it should be noted that removing that option increases the possibility that subjects will employ a strategy of selecting the easy items for immediate restudy simply to ''get them over with.'' (Son, 2004). We hope to mitigate that possibility by using somewhat shorter study lists (our lists use 48 pairs, whereas Son's (2004) used 60) and, in Experiments 2 and 3, by eliminating the JOL procedure. This procedural change reduced the total experimental time considerably.

The third change was to use items for which difficulty was manipulated experimentally. In Son's (2004) experiment, item difficulty was assessed via JOL. In Experiment 1, we manipulate difficulty and collect JOLs and thus have both an check of our manipulation and greater experimental control (because of random assignment of items to difficulty condition).

A final change was motivated by the findings of Metcalfe and Kornell (2003), who showed that the typical spacing effects reverse with very short presentation times. Thus, in Experiments 1 and 2, we chose to use longer presentation times (5 s for each presentation) than did Son (2004), who used 1 and 3 s study times (for first and second presentation, respectively). That is, we selected conditions under which we knew that a healthy spacing effect would obtain (in the first two experiments).

The present experiments were thus designed to elicit subjects' metacognitive control strategies for the massing and spacing of items under the experimental constraint that each subject must choose to mass exactly half of the items and space the other half, and could not choose to bypass a second presentation of any item. Study time for each item and for the entire study list were held constant. The study list was constructed so that half of the to-be-learned items were normatively easy and the other half were normatively difficult, although, as a manipulation check, JOLs were also solicited.

Experiment 1

Method

Subjects

Thirty-five students from an upper-level undergraduate psychology class at the University of Illinois at Urbana-Champaign participated in this experiment.

Design

Normative item difficulty (easy or difficult) and time between presentations of stimuli (massing or spacing) were manipulated within subjects. Dependent variables were judged difficulty (JOLs), spacing selections, and cued recall performance.

Materials

All words were 4–8 letter nouns. "Easy" pairs were obtained from the University of South Florida Free Association Norms (Nelson, McEvoy, & Schreiber, 1998). Only pairs with modest cue-target strength (forward strength between 0.01 and 0.03, inclusive) were used in order to minimize the chances of participants' guessing the correct target. "Difficult" pairs were generated by taking cues from the above normed list of nouns and pairing them randomly with nouns from another experiment. These difficult cue-target pairs were cross-checked with the USF norms to ensure each cue and target were not significantly associated. The difference between easy and difficult in the present study, then, is operationalized as the difference in the ease of learning paired associates versus unrelated word pairs.

The study list consisted of 48 randomly selected word pairs, 24 of which were easy pairs and 24 of which were difficult pairs. All presentation of stimuli and recording of subjects' responses were done on PC microcomputers programmed in QBASIC.

Procedure

Subjects worked individually in a large computer laboratory. Prior to study, subjects read a set of instructions informing them that they would be studying a list of word pairs, with each pair appearing exactly two times throughout the list, and that there would be a later test in which the word on the left (the cue) would be presented and the task would be to recall the appropriate target word. Subjects were instructed that they would first assess the difficulty of each pair and then choose whether they wanted to see it again "sooner" or "later," with the constraint that they must select "sooner" for exactly half (24) of the items and "later" for the other half. A running counter remained on the screen to remind them of their selection totals. The instructions suggested that, because of this restriction, they should "use a good strategy in deciding when to restudy the items," but no hints as to what constitutes a good strategy were given.

During the study phase, each word pair was presented for 5 s. At offset of the first presentation, the 1-to-6 difficulty scale appeared, and once this assessment was made, the prompt "Study 'S'ooner or 'L'ater" appeared, and subjects pressed the 'S' or 'L' key accordingly. Both the JOLs and the spacing decisions were self-paced. After a one-second interstimulus interval, the next item appeared.

The second presentation of each word pair was also 5 s in duration. A "massed" item's second presentation came after one intervening presentation of another word pair. Second presentations of all "spaced" items appeared in sequence after all first presentations and all second presentations of massed items. Thus, the last 24 studied word pairs were the second presentations of the spaced items (thereby intentionally confounding spacing with recency).

After subjects completed the study list, and were given a 5-min distractor task of simple math problems, they completed the self-paced cued recall test. Presentation of items in the test phase was done randomly. A cue appeared on the screen, with a blinking prompt where the target word had previously appeared. After a response was typed and entered, the next cue and blank appeared. A response was considered correct only when the entire word was spelled correctly.

Results and discussion

The results of all inferential statistics reported in this paper are either presented as difference scores with 1/2-width 95% confidence intervals (for pairwise comparisons) or as F tests (for interactions). All comparisons are reliable at the $\alpha = .05$ level using two-tailed tests unless otherwise noted. For the strategy selection data, in which accurate parameter estimation is the goal, all results are plotted with appropriate confidence intervals.

Judged versus normed difficulty

Proportions of easy and difficult items to which different JOL levels were assigned are shown in Fig. 1. Subjects assigned higher JOLs to easy than difficult items ($M_{\text{diff}} = 1.37$, 95% $\text{CI}_{\text{diff}} = \pm 0.16$), thus confirming the effectiveness of the relatedness manipulation.

Massing and spacing selections

Subjects chose more often to space the normatively difficult items, as shown in the rightmost bars of Fig. 2. That effect owed mostly to differences in selection in the second half of the experiment (that is, for the latter 24 items), as can be seen in the left and middle portions of Fig. 2.[1]

The same effect obtained when analyzing perceived difficulty and spacing choice, as shown in Fig. 3. JOLs of spaced items were lower than JOLs of massed items ($M_{\text{diff}} = 0.57$; 95% $\text{CI}_{\text{diff}} = \pm 0.43$).

Cued recall performance

As expected, recall of spaced items ($M = 0.81$) was higher than of massed items ($M = 0.68$; $\text{CI}_{\text{diff}} = 0.04$). Similarly, recall of easy targets ($M = 0.87$) was higher than of difficult targets ($M = 0.63$; $\text{CI}_{\text{diff}} = 0.05$). There was also a significant interaction, $F(1, 34) = 17.50$, suggesting that the spacing effect helped subjects' memory

[1] Note that, even though spacing and massing choices were constrained across the entire experiment, they were not constrained within each half of the experiment. Thus Fig. 2 shows separate bars representating the mean proportions of massed-easy items and spaced-difficult items. Across the entire list, those values are constrained to be equal, as shown by the bars of equivalent height on the right of the figure.

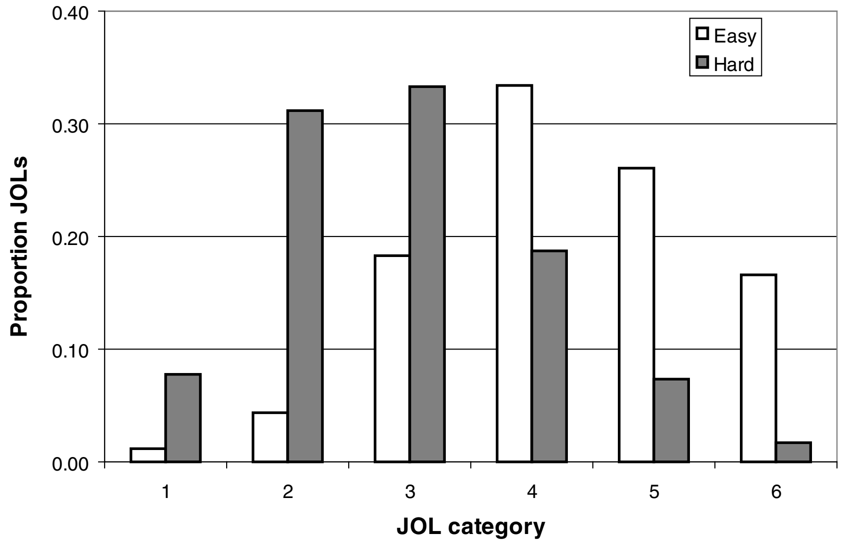

Fig. 1. Mean proportion of JOLs allocated to easy and difficult word pairs (Experiment 1).

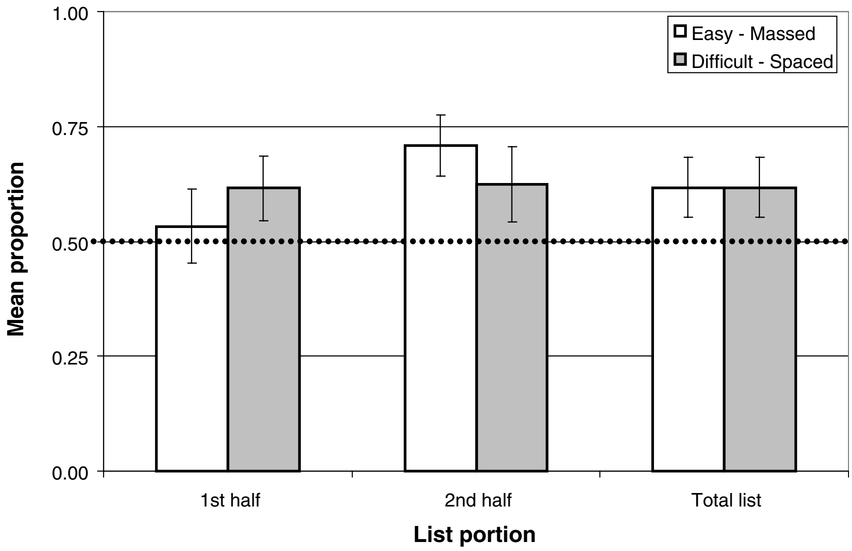

Fig. 2. Mean proportion of easy pairs chosen to be massed and difficult pairs chosen to be spaced as a function of list portion. Error bars represent the 95% confidence interval for the estimate. The dotted line indicates chance (i.e., no selective allocation across item types; Experiment 1).

for difficult items more than for easy items in the current paradigm, but overall differences in scale range renders any such conclusion premature. A breakdown of recall performance by JOL level and spacing choice can be seen in Fig. 4, which clearly displays the interaction between difficulty and spacing.

An additional analysis was conducted to investigate the association between strategy and ultimate performance on the test. Subjects at or above the median of number of spaced difficult pairs recalled a slightly higher proportion of items ($M = 0.77$) than subjects below the median ($M = 0.71$). This small effect will be considered later in the paper.

The reliable tendency of Ss to space difficult material (and thus mass easier items) is in contrast to the earlier results reported by Son (2004). We thus sought to replicate this effect and evaluate its efficacy in comparison with a control group that did not have control of the spacing of their study. In addition, because items' normed difficulty and judged difficulty were so highly related in Experiment 1 (as can be seen in Fig. 1), the JOL measure was eliminated.

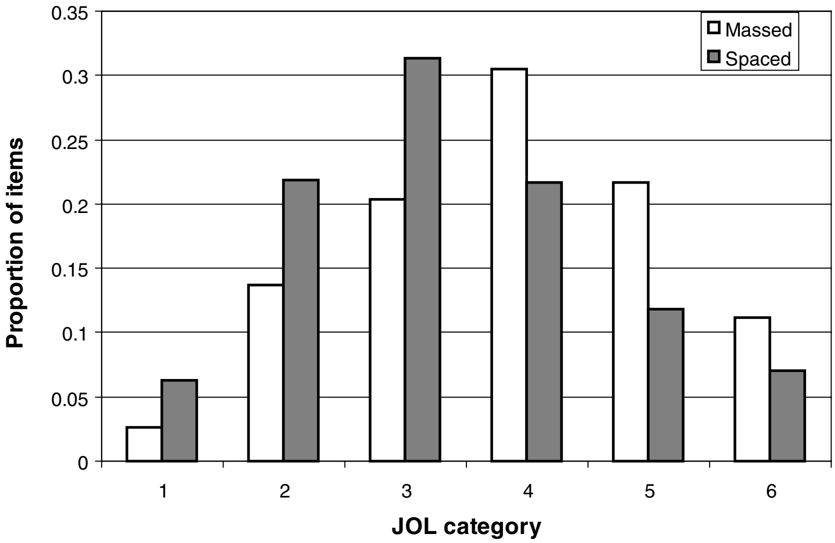

Fig. 3. Mean proportion of items selected to be massed or spaced as a function of JOL level (Experiment 1).

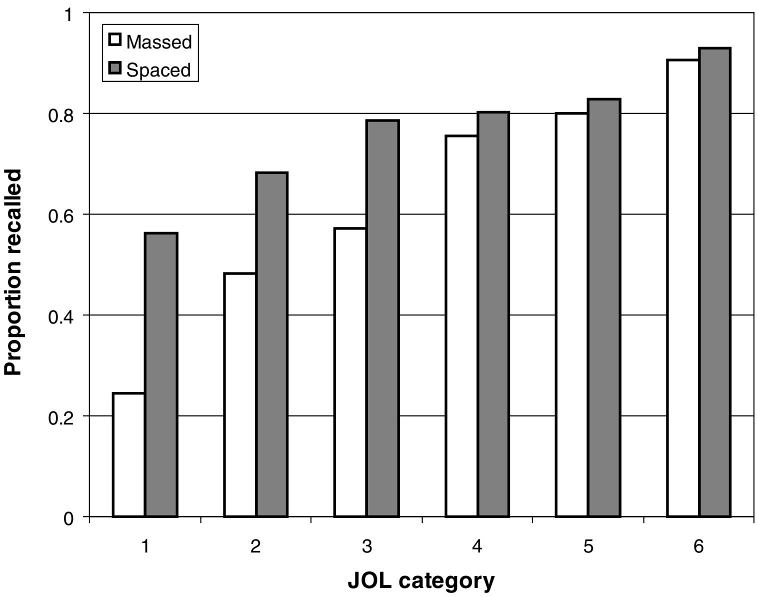

Fig. 4. Mean proportion recalled as a function of JOL and scheduling selection (Experiment 1).

Experiment 2

Method

Subjects

Eighty-eight subjects from an introductory psychology class at the University of Illinois at Urbana-Champaign participated as part of a course requirement. These subjects, unlike those in Experiment 1, had typically not had any previous instruction in cognitive psychology.

Design

Control of spacing (learner vs. experimenter) was manipulated between subjects. Item difficulty (easy vs. difficult) and time between presentations (massed vs. spaced) were within-subjects variables. Dependent variables were spacing selections and cued recall performance.

Materials

Materials were identical to those used in Experiment 1.

Procedure

Although most procedural details were identical to those used in Experiment 1, there were some important differences. Subjects were tested individually in a small laboratory. The instructions made no reference to judgments of difficulty, as none were to be collected. Thus, during the course of the experiment, at offset of the first presentation of a word pair, the spacing prompt appeared in lieu of the JOL prompt.

Subjects in the *self-control* condition retained complete control over the spacing of their study list, under the same restrictions as subjects were under in Experiment 1. Subjects in the *experimenter-control* condition were still asked to make item-by-item massing and spacing choices, but their inputs had no bearing on which items were actually massed or spaced. Instead, prior to beginning the study list, the computer program randomly selected half (24) of the study items to mass, and the other half to space. Subjects in both groups were informed that "[y]our decisions to re-study soon or later will have some, but not a total impact on when the computer re-presents the pairs."

Results and discussion

Massing and spacing selections

As in Experiment 1, subjects chose to space difficult items more often than easy ones. This effect in shown in Fig. 5, and also shows that, like Experiment 1, this effect appeared primarily in the second half of the list. Subjects in both the self-controlled ($M = 0.56$; 95% CI = ±0.06) and the experimenter-controlled ($M = 0.59$; 95% CI = ±0.06) conditions showed this effect.

Cued recall performance

Overall performance was almost identical for subjects in the self-control condition ($M = .60$) and those in the experimenter-control condition ($M = .61$). Thus, having the ability to choose when to re-study items did not benefit learners in this paradigm.

Analyzing test performance by item type for both groups revealed the expected advantage for recall of easy items ($M = 0.74$) over hard items ($M = 0.48$; $CI_{diff} = 0.04$), and spaced ($M = 0.69$) over massed items ($M = 0.52$; $CI_{diff} = 0.05$). The interaction between difficulty and spacing found in Experiment 1 did not obtain in Experiment 2 for subjects who controlled their own study, $F(1,43) = 1.77$, *ns*, nor subjects for whom spacing was assigned at random to easy and difficult items, $F(1,43) = .08$, *ns*. Thus, it appears that the form of the interaction present in Experiment 1—in which subjects made JOLs during study and consequently had higher overall performance scores—is more a function of scale location than a meaningful effect.

As in Experiment 1, subjects in the self-control condition who were at or above the median in terms of number of spaced difficult pairs recalled a slightly higher proportion of items ($M = 0.64$) than did subjects below the median ($M = 0.56$). Although this result might be taken to suggest the superiority of such a selection strategy, and thus stand in contrast to the results comparing conditions, it should be noted that subjects in the experimenter-control condition also showed this effect

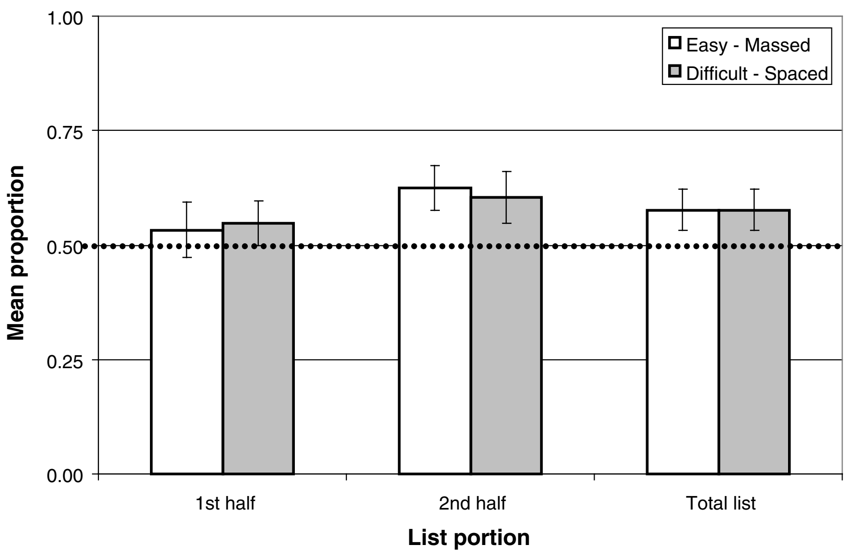

Fig. 5. Mean proportion of easy pairs chosen to be massed and difficult pairs chosen to be spaced as a function of list portion. Error bars represent the 95% confidence interval for the estimate. The dotted line indicates chance (i.e., no selective allocation across item types; Experiment 2).

($Ms = 0.65, 0.56$). Taken together, these results indicate that subjects with superior memory do indeed tend to space more of the difficult pairs, but the results do not support the idea that this strategy underlies that advantage.

The evidence that subjects mostly chose to temporally space their study of difficult material and mass the study of easier material is of consequence, and remains in contrast to earlier reports that subjects choose the opposite strategy (Son, 2004). Two critical differences between our procedure and that of Son (2004) are per-item study time, and exactly how massed the "massed" condition was. In her experiment, study times were much shorter than the present ones and the re-presentation of a massed item came immediately (unlike ours, in which there was one intervening pair). We reviewed findings earlier that suggested that massing can actually lead to superior recall performance under conditions more similar to Son's (Metcalfe & Kornell, 2003). The second, immediate re-presentation of a poorly encoded stimulus boosts the level of learning such that it is easier to recall that stimulus than it would be had it been presented rapidly at two spaced times during study. In the present experiments, each word pair's initial (and second) study trial was 5 s in duration, which was likely long enough to make subjects feel strongly about their learning of that item and want to "get it over with," such that the beneficial effects of spacing could be saved for more difficult items (Son, 2004). If this difference in presentation time and massing implementation between our experiments and Son's could explain the conflicting results, then eliminating those differences might cause subjects to shift their metacognitive study strategies. Thus, in Experiment 3, study time for each item was shortened, and items selected for massing were re-presented immediately.

Experiment 3

Method

Subjects

Thirty-six subjects from an introductory psychology class at the University of Illinois at Urbana-Champaign participated to fulfill part of a course requirement.

Design

Item difficulty (easy vs. difficult) and time between presentations (massed vs. spaced) were manipulated as within-subjects variables. Dependent variables were spacing selections and cued recall performance.

Materials

Materials were identical to those used in Experiments 1 and 2.

Procedure

The procedure was similar to that used in Experiment 1 with the following exceptions: Subjects were tested individually in a small laboratory. Each presentation was 0.5 s in length, with a 0.1 s interstimulus interval, meaning that each pair was on the screen for a total of only one second during the entire study phase. In addition, the second presentation of all massed items came immediately after the first, rather than having one intervening item between presentations.

Results and discussion

Massing and spacing selections

Under conditions in which presentation times were greatly reduced, subjects did not reliably choose to space difficult items, as shown in the right bars of Fig. 6. As in the previous experiments, however, there was a shift across test halves towards that effect (more specifically, to increase the allocation of massing decisions to easy items). Thus, although the effect evident in the previous two experiments was largely eliminated, there was no strong evidence that subjects chose to mass the difficult items, as was predicted.

Cued recall performance

As expected, shorter study times led to poorer recall than in previous experiments ($M = 0.32$). Recall performance was higher for easy ($M = 0.48$) than difficult items ($M = 0.15$; $CI_{diff} = 0.04$). Consistent with the findings of Metcalfe and Kornell (2003), there was no reliable advantage of spaced ($M = 0.33$) over massed study ($M = 0.30$; $CI_{diff} = 0.05$). The interaction between difficulty and spacing was also absent, $F(1,35) = .45$, *ns*. As in Experiments 1 and 2, there was a slight advantage in recall for subjects who were at or above the median ($M = 0.35$) in terms of number of difficult items chosen to be spaced over those who were below it ($M = 0.27$).

The important result from this experiment is that shortened study times—and the presumably lower level of learning for items that resulted—did not compel subjects to adopt a mass-difficult strategy. That is, even when learning of an easy item may not have plateaued and immediate additional study could prove fruitful, subjects did not choose to immediately restudy that word pair. Thus, it seems unlikely that the entire discrepancy between the space-difficult strategies reported in the present paper and the space-easy strategy reported by Son (2004) can be explained by differences in study time. Nonetheless, the reduction of the effect apparent in the first two experiments suggests that study time may be a factor.

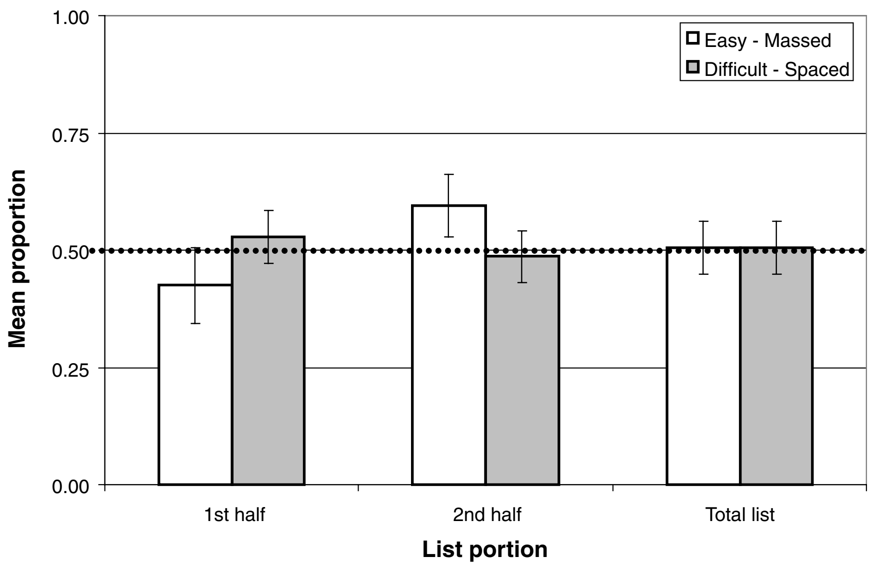

Fig. 6. Mean proportion of easy pairs chosen to be massed and difficult pairs chosen to be spaced as a function of list portion. Error bars represent the 95% confidence interval for the estimate. The dotted line indicates chance (i.e., no selective allocation across item types; Experiment 3).

General discussion

The primary purpose in conducting these experiments was to determine how subjects chose to space their study trials. We found that subjects preferred to space presentations of the difficult items and mass study of the easy items, and that this effect emerged more dramatically later in the list. This effect may reflect learning across the experiment or perhaps the greater immediacy of the spacing constraints when the choice limits are being approached.

Most importantly, this result conflicts with the finding reported by Son (2004) that subjects prefer to mass study of difficult material. Although there were numerous methodological differences between the studies, our results follow quite naturally from the discrepancy-reduction hypothesis, and thus salvage that viewpoint as a contributing mechanism to strategy selection in the scheduling of learning events. Reducing per-item study time was expected to elicit a different tactic from subjects by virtue of the reversal of the typical spacing advantage under very short study times (Metcalfe & Kornell, 2003). While this manipulation did eliminate subjects' preference to space difficult items, it did not result in the emergence of a space-easy strategy.

It is also possible that the knowledge that their massing and spacing options were limited caused subjects to use a stricter strategy than they would have if they could have massed or spaced as many items as they so desired. Analogous results have been obtained in study-time allocation experiments, and such constraints eliminate additional sources of variance that might work against detection of such an effect.

Study strategy optimality

Although the chief purpose of the present experiments was to establish subjects' metacognitive study strategies, the question of whether a particular strategy ultimately leads to better performance remains unresolved. Data from Experiment 1 showed that difficult items benefited more from spacing, but scaling concerns, as well as the fact that this result did not replicate, make any conclusions premature.

In fact, for subjects in Experiment 2, implementing one's own metacognitive study plan yielded no better performance than having massing and spacing assigned randomly to items. While one conclusion is that subjects chose to implement a suboptimal strategy, another possibility is that the effects of spacing one type of item and massing another type perfectly cancel each other out—at least with the stimuli and conditions used in these experiments—such that there *is* no optimal difficulty-driven strategy. The answer to this question may ultimately be task- or experiment-specific: If the "easy" items are so easy that only one trial is needed for mastery, the obvious optimal strategy would be to save the spacing benefit for difficult items. Conversely, if "difficult" items are so difficult that even spaced trials fail to lead to recall performance superior to massed trials, the obvious optimal strategy would be to use the spacing benefit on easy items. Thus, depending on the range of performance, multiple interactions might be possible; an intriguing possibility is that subjects' future choices about spacing might be occasionally misled into inappropriate generalizations by such spurious factors.

Implementing multiple study-test cycles (cf. Benjamin, 2003; Diaz & Benjamin, 2006) so that subjects may experience the effects of massing and spacing items of varying difficulty on performance may ultimately lead some subjects to change their metacognitive study strategies. However, the results in the present experiments showed no objective superiority of a space-difficult over a space-easy strategy, or vice versa. In the present paradigm, therefore, altering one's strategy could not be interpreted as an improvement in the use of one's metacognitive knowledge.

There is evidence, however, that subjects who have superior memory are more likely to show the tendency to space difficult rather than easy pairs. Across the three experiments, the correlation between number of difficult pairs chosen to be spaced and total recall is 0.19 ($t[157] = 2.43$). Although there is no evidence in these data that this strategy is useful, it must be considered that these experiments were simply inappropriate to reveal the usual superiority of such a strategy. Because this effect obtained both in subjects who had control over their spacing regimens and those who did not (in Experiment 2), it appears as though this effect reflects a preference of subjects with superior memory, but not a contributor to superior performance.

Future research must examine what beliefs a subject has about massing and spacing when investigating metacognitive spacing strategies. Although evidence suggests that subjects overestimate the advantages of massing and underestimate the effects of spacing (e.g., Simon & Bjork, 2001; Zechmeister & Shaughnessy, 1980), and generally *prefer* massing (Baddeley & Longman, 1978), it is not altogether clear that this preference does not obtain for other reasons—for example, because massing is less strenuous or because it leads to better immediate performance. The present data suggest that at least some subjects choose to selectively apply the advantageous effects of spacing to difficult material, which suggests that subjects' conceptions about spacing and massing may be more sophisticated than suggested by previous work.

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. (2003). Predicting and postdicting the effects of word frequency on memory. *Memory & Cognition, 31*, 297–305.

Bjork, R. A. (1994). Memory and metamemory considerations in the training of human beings. In J. Metcalfe & A. P. Shimamura (Eds.), *Metacognition: Knowing about knowing* (pp. 185–205). Cambridge, MA: MIT Press.

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.

Diaz, M., Benjamin, A.S. 2006. Effects of proactive interference (PI) and release from PI on judgments of learning. Manuscript under review.

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. C. Graesser (Eds.), *Metacognition in educational theory and practice* (pp. 249–275). Mahwah, NJ: Erlbaum.

Glenberg, A. M. (1977). Influences of retrieval processes on the spacing effect in free recall. *Journal of Experimental Psychology: Human Learning and Memory, 3*, 282–294.

Koriat, A., & Goldsmith, M. (1996). Monitoring and control processes in the strategic regulation of memory accuracy. *Psychological Review, 103*, 490–517.

Le Ny, J. F., Denhiere, G., & Le Taillanter, D. (1972). Regulation of study-time and interstimulus similarity in self-paced learning conditions. *Acta Psychologica, 36*, 280–289.

Mazzoni, G., Cornoldi, C., & Marchitelli, G. (1990). Do memorability ratings affect study-time allocation? *Memory & Cognition, 18*, 196–204.

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.

Nelson, D.L., McEvoy, C.L., Schreiber, T.A. 1998. The University of South Florida word association, rhyme, and word fragment norms. <http://www.usf.edu/FreeAssociation/>.

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.

Schwartz, B. L. (1994). Sources of information in metamemory: Judgments of learning and feelings of knowing. *Psychonomic Bulletin & Review, 1*, 357–375.

Shaughnessy, J. J. (1976). Persistence of the spacing effect in free recall under varying incidental learning conditions. *Memory & Cognition, 4*, 369–377.

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

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.

Young, C. J. (2004). Contributions of metaknowledge to retrieval of natural categories in semantic memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 4*, 909–916.

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.