ON THE EFFECTIVENESS OF SELF-PACED LEARNING

BY

JONATHAN G. TULLIS

THESIS

Submitted in partial fulfillment of the requirements
for the degree of Master of Arts in Psychology
in the Graduate College of the
University of Illinois at Urbana-Champaign, 2010

Urbana, Illinois

Master's Committee:

Professor Aaron Benjamin, Chair
Professor Brian H. Ross

ABSTRACT

Metacognitive monitoring and control must be accurate and efficient in order to allow self-guided learners to improve their performance. Yet few examples exist in which allowing learners to control learning produces higher levels of performance than restricting learners' control. Here we investigate the consequences of allowing learners to self-pace study of a list of words on later recognition, and show that learners with control of study-time allocation significantly outperformed subjects with no control, even when the total study time was equated between groups (Experiments 1 and 2). The self-pacing group also outperformed a group for which study time was automatically allocated as a function of normative item difficulty (Experiment 2). The advantage of self-pacing was apparent only in subjects who utilized a discrepancy reduction strategy—that is, who allocated more study time to normatively difficult items. Self-pacing can improve memory performance, but only when appropriate allocation strategies are used.

TABLE OF CONTENTS

INTRODUCTION

Being an effective student requires the ability to guide one's own learning activities effectively (e.g., Finley, Tullis, & Benjamin, 2009; Metcalfe, 2009). Students choose what to study, how to study, when to study, and how long to study; these self-regulated aspects of learning have important implications for the effectiveness of their learning efforts and achievement in education (Dunlosky & Theide, 1998). Recent research on the effectiveness of self-controlled study has revealed beneficial effects of allowing subjects choice about which items to re-study (Kornell & Metcalfe, 2006; Nelson, Dunlosky, Graf, & Narens, 1994) and how to schedule re-study (Benjamin & Bird, 2006; Son, 2004; Toppino, Cohen, Davis, & Moors, 2009). Here we consider the effectiveness of allowing learners control over the allocation of study time across a heterogeneously difficult set of items. In addition, we investigate how individual differences in allocation policy affect the benefits of according control to learners.

In the experiments presented here, we directly compare the effects of self-control over study time with control conditions in which total study time is equated and items are allocated either equivalent study time (Experiments 1 and 2) or are allocated study time based on their normative difficulty (Experiment 2). As we review shortly, the extant literature on control over study time provides no clear answer on whether such control is beneficial and even lesser guidance on the control strategies that affect those potential benefits.

The effect of control of study time allocation is related to two metacognitive skills that have been studied in the metacognition literature. First, learners must be able to apply an effective learning strategy to a heterogeneously difficult set of items. Second, monitoring of the items must be reasonably accurate—that is, subjects must be able to discern which items will be easy and which will be difficult for them to learn. We treat these issues in turn.

Strategies in allocation of study time

One of the most fundamental aspects of self-regulated learning is deciding how to allocate study time across items. Learners usually devote more time to items they judge to be difficult, at least when no time constraints exist. A meta-analysis by Son and Metcalfe (2000) revealed that, out of 46 published conditions examining the relationship between difficulty and study time, 35 revealed a preference to devote more study time to material judged to be difficult. Individual differences in the degree to which subjects modulate their study time across items have also been noted. Both younger and older adults spend more time studying difficult items, but younger adults modulate their study time based upon item difficulty to a greater extent than do older adults (Dunlosky & Connor, 1997).

Two general theoretical approaches explain how subjects allocate study time by outlining the decision process used to cease study of an item. The discrepancy reduction theory suggests that people stop studying an item once that item's judged level of learning meets a pre-set criterion (Dunlosky & Theide, 1998), while the region of proximal learning theory (Metcalfe & Kornell, 2003) suggests that people stop studying an item once the "rate of return" (amount of mnemonic benefit per time unit of study) falls below a criterion rate. Although they rely upon different cessation-of-study mechanisms, discrepancy reduction and the region of proximal learning theories often predict the same self-paced study behavior: people study the difficult items longer than the easy ones.

Memory Monitoring

Both theoretical approaches assert that learners base self-regulation directly on metacognitive monitoring, a view that is supported by evidence that *judgments of learning* (JOLs) are dissociable from recall (e.g., Benjamin, 2003; Benjamin, Bjork, & Schwartz, 1998;

Metcalfe, Schwartz, & Joaquim, 1993) but nonetheless correlated with later study choices (Metcalfe & Finn, 2008). Because study choices appear to be directly based upon introspective monitoring of memory (Nelson & Narens, 1994; Metcalfe & Finn, 2008), the effectiveness of self-regulation faces a major bottleneck in the accuracy of memory monitoring. Giving learners control over study allows biases and inaccuracies in metacognitive monitoring to influence performance negatively by leading learners to make suboptimal or counterproductive decisions while controlling their study.

Memory assessments are prone to inaccuracies, sometimes dramatically so. For example, poor students often overestimate their performance on tests by up to an astonishing 30% (Dunning, Johnson, Ehrlinger, & Kruger, 2003; Hacker, Bol, Horgan, & Rakow, 2000). An insidious problem is ineffective monitoring across materials within a task. Learners who exhibit poor relative accuracy (or resolution) are not able to distinguish between materials that will be relatively easy to learn and ones that will be challenging. Judgments concerning levels of learning and states of knowledge exhibit poor resolution under many conditions (Dunlosky & Nelson, 1992; Dunlosky & Nelson, 1994; Shaughnessy, 1981), and often do not appropriately reflect the influence of variables that affect performance considerably (Benjamin, 2005; Koriat, Bjork, Sheffer, & Bar, 2004; Karpicke & Roediger, 2008; Zechmeister & Shaughnessy, 1980, Roediger & Karpicke, 2006).

The sophistication of learners' mental models and memory monitoring ability determines the potential effectiveness of self-guided learning. Whether the opportunity for control plays out in a net gain for self-controlled processing may depend on individual differences in metacognitive skill and the difficulty of the metacognitive aspects of the task. For example, learners who are afforded the opportunity for accurate metacognitive monitoring (by delaying the occasion of their JOL until well after the initial exposure) choose to re-study less well learned materials than learners who have no such diagnostic monitoring opportunity. The differences in accuracy of metacognitive monitoring lead to differences in re-study choices, which lead to differences in performance (Thiede, Anderson, & Therriault, 2003).

Efficiency of self-paced study

Although there is much research concerning how subjects allocate study time across different materials (Baker & Anderson, 1982; Maki & Serra, 1992), with different criterion tasks **(**Le Ny, Denheire, & Taillanter, 1972; Zacks, 1969; Belmont & Butterfield, 1971**)**, and even with varying subject populations (Dufresne & Kobasigawa, 1988, 1989; Kobasigawa & Metcalf-Haggert,1993, Dunlosky & Connor, 1997), the basic question of whether people are effective at distributing their own study time compared to a group that does not control their own time distribution remains unresolved. Metcalfe and Kornell (2005) noted that, although there are ample data on <u>how</u> learners allocate study across items, "we still do not know whether what they do enhances their learning, or is in any way optimal."

Only two studies (Mazzoni & Cornoldi, 1993; Koriat, Ma'ayan, & Nussinson, 2006) have addressed whether people can effectively allocate their study time. Mazzoni and Cornoldi (1993) compared a group that was allowed to self-pace study with a control group that viewed words presented at a constant rate. The constant rate for subjects in that control group was determined by the average rate with which those subjects chose to study items in a previous list. The self-paced group recalled more words than the control subjects; however, Koriat, Ma'ayan, and Nussinson (2006) did not replicate this result when they compared a self-pacing group to a control group that was equated on total study time. In their experiment, each control subject was yoked to a self-pacing subject in terms of word order and total study time, but the study time was divided evenly between the words. With these differing results, it is difficult to conclude that allowing learners to self-pace actually benefits their overall performance. In this paper, we evaluate whether self-pacing is efficacious even when a strict total-study-time control is implemented.

Given the many conditions in which metacognitive monitoring is inaccurate, and the numerous opportunities for subjects to implement an ineffective control policy, it is not obvious that subjects can effectively allocate study time to improve their memory. In addition, it has been theorized that spending more time on difficult items may be unwise, since there are conditions under which those items provide a low rate of return (Metcalfe & Kornell, 2003). By spending more time on the difficult items, self-pacing subjects will necessarily devote a smaller proportion of time to the easy- and medium-difficulty items, and this choice could adversely impact overall performance. In Experiment 1, we compare the memory of subjects who choose how to allocate their study time with the memory of subjects who spend the same total time as the self-pacers, but have no control over their study schedule. Control subjects viewed the same words as self-paced subjects, but viewed them for the average amount of time taken per word by their self-paced partner.

EXPERIMENT 1

We use a procedure similar to that of Koriat et al. (2006), in which control subjects were yoked to self-paced subjects in total study time. Whereas previous studies have utilized free recall or cued recall to measure memory, we utilized recognition. There are numerous advantages to using recognition. First, because recognition testing affords fewer opportunities for test strategies to differ across subjects (Benjamin, 2008), it allows us to more precisely focus on the effects of strategies employed during encoding. Second, because recognition tasks are metacognitively deceptive for learners—that is, because the modal expectation of a memory test appears to conform more to the demands of recall than recognition testing (Benjamin, 2003; Glanzer & Bowles, 1976; Guttentag & Carroll, 1998)—finding evidence for the effectiveness of metacognitive control on recognition meets a somewhat higher standard and should generalize to other memory tasks. Finally, motivated by the inconsistent prior findings using free recall, we sought a test paradigm with high power and in which decision strategies can be effectively corrected for. Recognition meets these dual constraints because many more items can be used than in recall paradigms, and because we can bring the Theory of Signal Detection (Green & Swets, 1966; Macmillan & Creelman, 2005) to bear on eliminating individual differences in the decision component of the recognition task (but see Benjamin, Diaz, & Wee, 2009).

Method

Participants. One-hundred and forty eight introductory-level psychology students from the University of Illinois participated in exchange for course credit.

Materials. One-hundred and sixty words were selected from the MRC Psycholinguistic Database (Wilson, 1988), and ranged on measures of familiarity (range 214-657, mean = 495.00, SD = 111.52), concreteness (range 186-645, mean=444.83, SD = 163.18), and imageability (range: 210-667, mean= 467.48, SD = 142.15). For each pair of subjects, eighty of these words were randomly selected to comprise the study list. The selection and order of the presented words were randomized and the presentation list included four primacy and four recency buffer words that were excluded from analysis. All 160 words comprised the test list.

Design and Procedure. Subjects were run individually on desktop computers in individual rooms. The first subject on a given computer room was assigned to the self-paced condition. The next subject on that computer arrived after the first subject completed the experiment and was assigned to the fixed-rate condition. Thereafter, subjects were alternately assigned to self-paced and fixed-rate conditions. The first subject in each pair studied each word for as long as they chose and proceeded to the next word by pressing the space bar. The fixed-rate subjects viewed the same list of words in the same order as their yoked partners, but each word they viewed was presented for a constant amount of time, determined by calculating the average amount of time per word taken by the yoked subject. Words were presented individually in the middle of a white computer screen in black, Times New Roman, 60 point font. All subjects were instructed to "do your best to remember the words for a later memory test."

After viewing the target list of 80 words, subjects engaged in a recognition task that included the 80 studied items and 80 previously unstudied items. The order of the 160 words was randomly determined and each word remained on the screen until subjects provided a recognition judgment on a scale of 1 to 4. This scale ranged from "I am certain I have not seen that word" (1) to "I think I have not seen that word" (2) to "I think I have seen that word" (3) to "I am certain I have seen that word" (4). Each yoked pair of subjects received the test words in the same order. Confidence ratings were used to generate a measure of discrimination (d_a) between old and new items based on unequal-variance signal-detection theory (Green & Swets, 1966; for a discussion of the advantages of such a measure in recognition, see Matzen & Benjamin, 2009)

Results

Mean discriminability (d_a) is shown in the left panels of Figure 1 (subject analysis) and Figure 2 (item analysis), and hit and false alarm rates are shown in Table 1. All statistics reported here are significant at an $\alpha < .05$ level. Self-paced subjects revealed significantly higher performance than their paired controls ($d_{a[selfpaced]}$=1.73 [SEM=0.07], $d_{a[fixed\text{-}rate]}$=1.51 [SEM=0.07]; t [73] = 2.20[1]) despite being equated on total study time, revealing that the individual control of study time is beneficial for learning. A similar result is obtained following an item analysis, in which discrimination was calculated for each word once from the combined responses of the entire group of self-paced subjects and once from the combined responses of the entire group of control subjects. Item discriminability calculated from the self-paced subjects was significantly greater than that computed from the control subjects (d_a=1.58 [SEM=0.06], d_a=1.43 [SEM=0.05]; t [159]=4.25).

Benefits of self-pacing as a function of individual differences in control strategy

We mentioned previously that individual differences in how study time is allocated are likely to affect the outcome of experiments in which strategic control is experimentally manipulated. Here we consider study time allocation as a function of item difficulty and classify learners into two general categories. Learners who spent more time studying the more normatively difficult items were considered discrepancy reducers (Dunlosky & Thiede, 1998), and learners who allocated more time to easier items were considered to be discrepancy increasers.

[1] In a design like this one, one pair of yoked subjects is treated as a single unit of analysis. A traditional between-subjects comparison would not account for the pairwise yoking in the design.

To evaluate an individual subject's allocation policy, we calculated the correlation across items between allotted study time and normative item difficulty, d_a, as assessed by mean performance on that item by the fixed-rate control group. The average correlation between study time and normative word difficulty for all subjects in the self-paced condition was not significantly different than zero ($r = 0.02$, $SD= 0.13$; $t[73] = 1.33$). Thirty-five subjects exhibited negative correlations, classifying them as discrepancy reducers, and 39 subjects revealed positive correlations, classifying them as discrepancy increasers.

The performance of each of these subsets of subjects is compared to their yoked controls, and the mean levels are shown in the right panels of Figures 1 and 2. Discrepancy reducers exhibited reliably higher levels of discriminability than their yoked controls ($d_{a\ [reducers]}=1.77$ [SEM=0.10], $d_{a\ [yoked\ to\ reducers]}=1.42$ [SEM=0.10]; $t[34] = 2.41$), but discrepancy increasers did not ($d_{a\ [increasers]}=1.69$ [SEM=0.11], $d_{a\ [yoked\ to\ incressers]} =1.59$ [SEM=0.10]; $t[38] = 0.75$). An item analysis based on the same breakdown of subjects is shown in Figure 2, and replicates this finding ($d_{a[reducers]}=1.62$ [SEM=0.05], $d_{a[yoked\ to\ reducers]}=1.30$ [SEM = 0.05]; $t[159] = 6.04$; $d_{a[increasers]}=1.49$ [SEM=0.05], $da_{[yoked\ to\ increasers]}=1.47$ [SEM=0.05]; $t[159] = 0.38$, respectively). These results hint that allocation strategy modulates the benefits that self-pacing provides. A 2-way ANOVA on self-pacing condition (self-pacing or yoked) and allocation strategy (reducers or increasers) showed only a significant effect of self-pacing condition ($F(1,72)=5.20$) indicating that self-pacing benefits memory performance. Given that the ANOVA did not reveal a significant interaction between condition and strategy, the modulating effect of self-paced strategy is unclear; we will further investigate the impact of differing allocation strategies on performance in Experiment 2.

Discussion

Self-paced learning significantly improves memory performance compared to a control condition in which the same overall study time was used but time was not differentially allocated across items. By choosing how to distribute study time, subjects outperformed others who spent the same total time studying the material. The improvement seemed to be driven principally by the subjects who devoted more time to difficult words than to easy words.

Metacognition can be used effectively to improve cognition. However, only some learners utilize metacognition effectively as the strategy utilized during self-paced learning modulates the benefits that self-pacing provides. The act of self-pacing in and of itself does not improve memory; rather, self-pacing only yields improvements under certain circumstances. Some learners utilize suboptimal self-regulation techniques and perform no better than those who do not regulate their learning. Suboptimal metacognition may result from a lack of motivation or of knowledge; self-pacers may not have the motivation to monitor and control their learning or may not have the requisite knowledge to do so effectively. This motivation hypothesis will be contradicted, however, by evidence in the general discussion.

One concern in this experiment is that the difference between the yoked control groups is as large as the difference between the discrepancy reducers and increasers. Because subjects were not randomly assigned to the allocation variable, and consequently neither were the yoked controls, such an effect can reflect uncontrolled differences in study-list composition. One goal of Experiment 2 is to replicate Experiment 1 with sufficient power to directly compare discrepancy reducers and discrepancy increasers, thus evading concerns over the role of the control group in undergirding this individual-difference effect.

In addition, we evaluate whether the advantage of self-pacing arises only from the fact that normatively more difficult items are studied for greater amounts of time, or whether idiosyncratic differences in difficulty are also important. To test these hypotheses, Experiment 2 included a condition in which study times are automatically distributed based on normative item difficulty.

This contrast is relevant to educational and technological decisions concerning the best way of implementing study regimens. If it is the case that normative difficulty is an effective proxy for idiosyncratic difficulty, then automated study routines can bypass self-control—which may be unduly influenced by individual differences in motivation or other factors—in developing an effective study regimen. However, if intraindividual differences in item difficulty are high compared to intraitem differences in normative difficulty, then such a short-cut will not prove effective.

EXPERIMENT 2

It is unclear whether the opportunity to self-regulate in a word-learning task affords any advantages over a regimen in which normatively difficult items are allotted greater study time. There are strong reasons to suspect that such normative allotment of study time may be superior. First, the tacit individual assessments of item difficulty that precede study-time allocation are prone to the many biases and inadequacies discussed earlier, whereas normative data on item memorability are not susceptible to such sources of error. Second, control over self-pacing allows for individual differences in metacognitive sophistication to play a major role; when study times are allocated automatically, individual differences in allocation policy are irrelevant, and any beneficial effects will likely be evident for a larger subset of subjects. On the other hand, there is also one major reason to predict that normative allocation of study time will be inferior to self control. If idiosyncratic individual differences in item difficulty are larger than intra-list variance in normative item difficulty, normative assessments of that difficulty will be inadequate for a large subset of subjects.

We will evaluate these hypotheses in this experiment by forcing subjects in one condition to adopt a discrepancy reduction strategy based upon normative difficulty. We are interested in whether self-pacers' performance can be improved by tying their study time to item difficulty, such that everyone in this condition becomes a perfect discrepancy reducer. The materials and procedure for the second experiment were the same as those used in the first experiment with an additional condition in which presentation times were determined based on an algorithm relating normative discriminability (from the control group in Experiment 1) to study time.

Method

Participants. Two hundred and thirty-four introductory-level psychology students from the University of Illinois participated in exchange for course credit.

Design and Procedure. Subjects were assigned to groups of three based upon the order of their arrival; the first subject completing the experiment in each individual room was assigned to the self-paced condition, the second to the fixed-rate condition, and the third to the normative allotment condition. The latter two subjects were given the same total study time as taken by the self-paced subject in that group. The self-paced and fixed-rated conditions were the same as in Experiment 1, and the third condition provided differential study times based on normative difficulty. The presentation times were scaled in this group such that words of normatively lower discriminability were displayed for longer times than words of higher discriminability. We calculated the presentation times with the following formula:

$$ST_i = \frac{TT}{\sum_{allwords}\left(\frac{1}{d_a{}^v}\right)} \times \left(\frac{1}{d_{a,i}^v}\right),$$

in which ST_i is the study time for item i, TT is the total study time allotted to this subject, and d_a is the detection-based measure of memory discriminability for that item (based on performance by the control group in Experiment 1). The parameter *v* was manipulated slightly (v= 2.5 in the high variability group and v= 1.5 in the low variability group) across two different subgroups of subjects. The former value was chosen for half of the subjects in order to match the standard deviation of presentation times for the normative-allotment subjects to the standard deviation of study times that resulted from self-paced subjects in Experiment 1. However, when the large disadvantage in performance in this condition became evident, the value of v was reduced in order to decrease the variability in presentation times across items. It was a concern that the unpredictability in study time per item might be so jarring that performance would suffer, so we reduced the variability of presentation times for half of the subjects.

Results

Within the normative-allotment group, subjects' overall discriminability did not differ between the high and low variability groups ($d_{a[high]}$ =0.93 [SEM=0.07], $d_{a[low]}$=1.06 [SEM=0.08]; t[76] = 1.12), so subjects were combined across the levels of this variable. Performance in all three conditions is shown in Figures 1 and 2. A one-way ANOVA revealed a significant effect of condition (F(2,75)=11.16) . Post-hoc comparisons revealed that self-paced subjects showed higher discriminability than the control group ($d_{a[self\text{-}paced]}$=1.48, $d_{a[fixed\text{-}rate]}$=1.27, [SEM=0.11]; t[77] = 1.99; p=0.05), as in Experiment 1, and subjects in the normative-allotment condition showed lower discriminability (d_a=0.99) than both the self-paced ([SEM=0.11]; t[77]=4.41) and control ([SEM=0.10]; t[77] = 2.76) subjects. As shown in Figure 2, an item analysis obtained similar results, in which words were significantly easier to discriminate for self-paced subjects than for control ($d_{a[self\text{-}paced]}$=1.39 , $d_{a[fixed\text{-}rate]}$=1.21 [SEM=0.03]; t[159] = 4.21) or the normative-allotment (d_a=0.93, [SEM]=0.03; t[159]=10.59) subjects. Words were also significantly easier to discriminate for control subjects than for the normative-allotment subjects ([SEM=0.03], t[159]=8.49).

Benefits of self-pacing as a function of individual differences in control strategy. Self-paced subjects were again divided into two groups based on the correlation between study time and item difficulty as measured in the control group. In Experiment 2, the average correlation between item difficulty and study time allocation was again not different than zero (corr = 0.03, SD = 0.12, t(77)=1.78). Thirty out of 78 self-paced subjects exhibited a negative correlation between discriminability and self-paced study time and were therefore classified as discrepancy reducers. As shown in the right panels of Figures 1 and 2, and replicating what was seen in Experiment 1, only the discrepancy reducers exhibited better memory performance than their control subjects (discrepancy reducers: ($d_{a\,[self\text{-}paced]}$=1.77, $d_{a[fixed\text{-}rate]}$=1.17 [SEM=0.17], t[37]=3.47; discrepancy increasers: ($d_{a\,[self\text{-}paced]}$=1.30 $d_{a\,[fixed\text{-}rate]}$ =1.33 [SEM=0.13]; t[47]=0.22). In contrast to Experiment 1, a direct comparison between self-paced discrepancy reducers and self-paced discrepancy increasers showed a significant advantage for discrepancy reducers (t(77)=2.30). Importantly, neither the control nor normative-allotment subjects yoked to the discrepancy reducers differed from those yoked to discrepancy increasers in performance (t[77] = 1.01; t[77] = 0.72 respectively), indicating that differences in total study time and idiosyncratic item selection did not play a role in producing the advantage for subjects exerting self-control over study.

Discussion

The results of the Experiment 2 replicate those of Experiment 1: self-pacers exhibited superior recognition to control subjects, and this effect was driven by the subset of self-pacers who were discrepancy reducers. Because the benefits of discrepancy reducers are entirely correlational, however, it is impossible to conclude with any certainty that the strategy itself—rather than any correlated intellectual or motivational variables—fostered the performance advantage. These third variables will be considered more closely in the General Discussion. Experiment 2 also shows that subjects in the normative-allotment condition performed significantly worse than those in the other two conditions. This surprising result indicates that an allocation policy based on normative difficulty is not superior to one completely controlled by subjects and based upon their idiosyncratic metacognitive judgments.

GENERAL DISCUSSION

In both Experiments 1 and 2, recognition performance for learners who controlled their own study time was compared to learners who spent the same amount of overall study time but viewed the items for a standard amount of time. Learners with control of study-time allocation significantly outperformed subjects with no control, even when the total study time was equated between groups. This effect was driven only by those subjects who allocated their time in a manner consistent with a discrepancy-reduction strategy. In the second experiment, subjects in the third condition (normative-allotment subjects) viewed each item for a length of time determined by that item's objective difficulty. These learners performed significantly worse than both self-pacers and control subjects.

When data from the two experiments are combined, self-paced discrepancy reducers showed higher discriminability than self-paced discrepancy increasers ($d_{a[\text{reducrers}]}=1.77$, $d_{a[\text{increasers}]}=1.53$ $t(153)=2.14$), but the control subjects yoked to the reducers did not perform differently than control subjects yoked to the increasers ($d_{a[\text{yoked to reducers}]=}1.32$, $d_{a[\text{yoked to increasers}]}=1.45$ $t(153)=1.32$). Further, a 2-way ANOVA on self-pacing condition (self-paced or yoked) and allocation strategy (reducer or increaser) from the data combined across both experiments reveals both a benefit of self-pacing ($F(1,150)=11.88$) and an interaction between self-pacing condition and allocation strategy ($F(1,150)=9.30$). This result indicates that the strategy utilized during self-paced learning modulates the benefit of having control and allays our concern over the reliability of the individual-difference effect that was questionable in the Experiment 1 data alone.

Giving learners more control over their study behavior resulted in better memory performance, even without increasing total study time. Evidence presented here supports a growing trend to trust the metacognitive capabilities of learners and accordingly allow more self-regulation during learning (Finley, Tullis, & Benjamin, 2009; Kornell & Bjork, 2007; Kornell & Metcalfe, 2006). Although they must devote resources to the act of self-regulation, subjects with the opportunity to regulate study time still outperformed their yoked counterparts.

Control of one's own study, however, proves not inherently advantageous; specific strategies mediate the effectiveness of control. In this experiment, only discrepancy reducers outperformed their yoked partners and discrepancy reducers made up less than half of all the self-paced subjects tested. In addition to strategy utilized during learning, other individual differences exist in the effectiveness of implementing control. The degree to which the metacognitive strategy is implemented and the level of modulation utilized by the learner may influence the effectiveness of control. For example, older adults are shown to modulate their self-paced study time to a lesser extent than younger adults, and some suggest that older adults' deficient utilization of control during study contributes to their deficits in memory performance (Dunlosky & Connor, 1997).

Benefits of self-pacing

In order to more closely examine the advantages evidenced by self-paced subjects, we conducted an analysis of the relationship between normative item difficulty (estimated from subjects in the control condition of Experiments 1 and 2) and the difficulty experienced by self-paced subjects across both experiments. We have plotted item difficulty as calculated from self-paced subjects against item difficulty as calculated from control subjects in Figure 3. If self-pacing proves advantageous for items across the difficulty spectrum, then the function relating item difficulty between the groups will lie parallel to the major diagonal but have a higher intercept. If self-pacing affects item difficulty differently for easy than for difficult items, the slope of the line will significantly differ from the diagonal. Several important effects are evident in Figure 3. First, the intercept of this line is higher than 0, indicating that self-pacing reduced the overall difficulty of the items (intercept = 0.34; SE = 0.08; $t(157)=4.46$). Further, the slope of the line for self-paced subjects is shallower than the diagonal (slope=0.88; SE=0.05; $t(157)=2.14$), revealing that self-pacing reduced the heterogeneity of item difficulties. Self-pacing improved performance on difficult items to a greater extent than on the easier items. The position of this function reveals that self-pacers outperformed yoked, fixed-rate subjects by increasing performance for the normatively difficult items without sacrificing performance on the easier items.

Selection of allocation policy

Across both experiments, less than half of the self-pacers engaged in a discrepancy-reduction strategy. The failure to choose the more beneficial strategy could be caused by several things. Subjects may lack the motivation or knowledge to implement a good strategy, have low performance goals, or attempt to implement a discrepancy reduction strategy but fail to accurately monitor item difficulty.

Two pieces of evidence suggest that the differences between discrepancy increasers and reducers are not motivational. A comparison of the standard deviations of the presentation times that result from discrepancy reducers' and increasers' allocation behavior reveals no significant differences between the extent to which increasers and reducers regulate their learning (2.59 vs. 2.22; $t[151]=0.83$). If one assumes that lower motivation would translate into less variability in study time across items, then this result suggests that neither group was more motivated than the other. This result is difficult to interpret however, since it is a null result, and the power to detect a medium-sized effect (Cohen, 1988) is only 0.86. In addition, the time spent per word did not differ significantly between discrepancy increasers and reducers ($t_{[increaser]}$=3.1 s vs. $t_{[reducer]}$=3.8 s, $t[151]=1.68$, $p = 0.1$). Discrepancy reducers, then, do not appear to exhibit greater levels of motivation during study, as they spend similar amounts of time studying and modulate their learning to roughly the same degree as discrepancy increasers.

Previous research shows that subjects overwhelmingly choose to devote more time to items that they judge to be difficult (Son & Metcalfe, 2000), which suggests that the failure of subjects to follow a discrepancy reduction strategy may result from a lack of awareness of which items will be difficult at the time of the test or from low performance goals (Dunlosky & Thiede, 2004). Dunlosky and Thiede (2004) showed that when given significant time constraints or very low performance goals, subjects shift their study strategies from spending more time on difficult items to studying only the easy-to-master subset of materials. Discrepancy increasers in the current study may be subjects who adopted low performance goals and therefore devoted time disproportionately to the easier items. This line of reasoning suggests that discrepancy increasers would likely spend less time overall studying. However, the analysis presented previously does not support this claim. On the basis of this evidence, both the motivational and low performance explanations seem unlikely; the most viable explanations remaining are that subjects are discrepancy increasers because they cannot predict individual items' test difficulty during study or because they employ a bad metacognitive strategy.

Benefits of self-regulation of study time

Learning can be enhanced through successful implementation of self-guided study time allocation. But, as Metcalfe (2009) argues, there are two necessary components for control of study to be helpful: monitoring must be accurate, and appropriate choices must be implemented during study. Without accurate monitoring, appropriate choices can neither be made nor implemented. In this study, less than half of all subjects implemented an effective study strategy. Improvements in metacognition could result in more successful and efficient learning, both in and out of the classroom. Flavell (1979) offers a broad, hopeful vision for the usefulness of metacognitive instruction, suggesting that metacognitive ideas “could be parlayed into a method of teaching children (and adults) to make wise and thoughtful life decisions as well as to comprehend and learn better in formal educational settings” (p.910). For individuals who have deficits in their ability to monitor or control study, direct instruction or experience in metacognitive monitoring and allocation strategies can improve student’s choices and learning (see Finley et. al., 2010). Information technology also plays a promising role in training metacognitive monitoring and control, with the hope that student learning will be enhanced (for examples see Finley et al., 2010). In these experiments, allowing learners control their own learning greatly boosted performance and control strategy utilized during self-regulated learning modulated performance. This research thus links metacognitive control with performance, showing that differences in how control is utilized correspond to differences in memory. The research further suggests that individual differences in the use of control are meaningful to ultimate learning and more research should thus be devoted to exploring these differences. Overall, trusting learners to control their own learning, even when that control is based upon their metacognitive biases and misconceptions, might actually improve cognition.

TABLES

Table 1. Percent hits and false alarms (and standard deviations) for Experiments 1 & 2.

	Self-paced		Fixed-rate		Normed	
	Hits	False Alarms	Hits	False Alarms	Hits	False Alarms
Exp. 1						
All	0.76 (0.14)	0.14 (0.10)	0.71 (0.15)	0.16 (0.11)		
Reducers	0.77 (0.13)	0.14 (0.07)	0.67 (0.17)	0.15 (0.08)		
Increasers	0.76 (0.14)	0.15 (0.13)	0.75 (0.13)	0.17 (0.13)		
Exp. 2						
All	0.74 (0.13)	0.18 (0.14)	0.68 (0.14)	0.20 (0.13)	0.62 (0.13)	0.23 (0.13)
Reducers	0.80 (0.10)	0.18 (0.13)	0.66 (0.14)	0.21 (0.12)	0.64 (0.13)	0.23 (0.12)
Increasers	0.70 (0.14)	0.18 (0.14)	0.69 (0.14)	0.19 (0.14)	0.60 (0.14)	0.23 (0.14)

FIGURES

Figure 1. Left panel: Item discriminability as a function of condition based on a subject analysis. Right panels: Discriminability for discrepancy reducers (top) and discrepancy increasers (bottom).

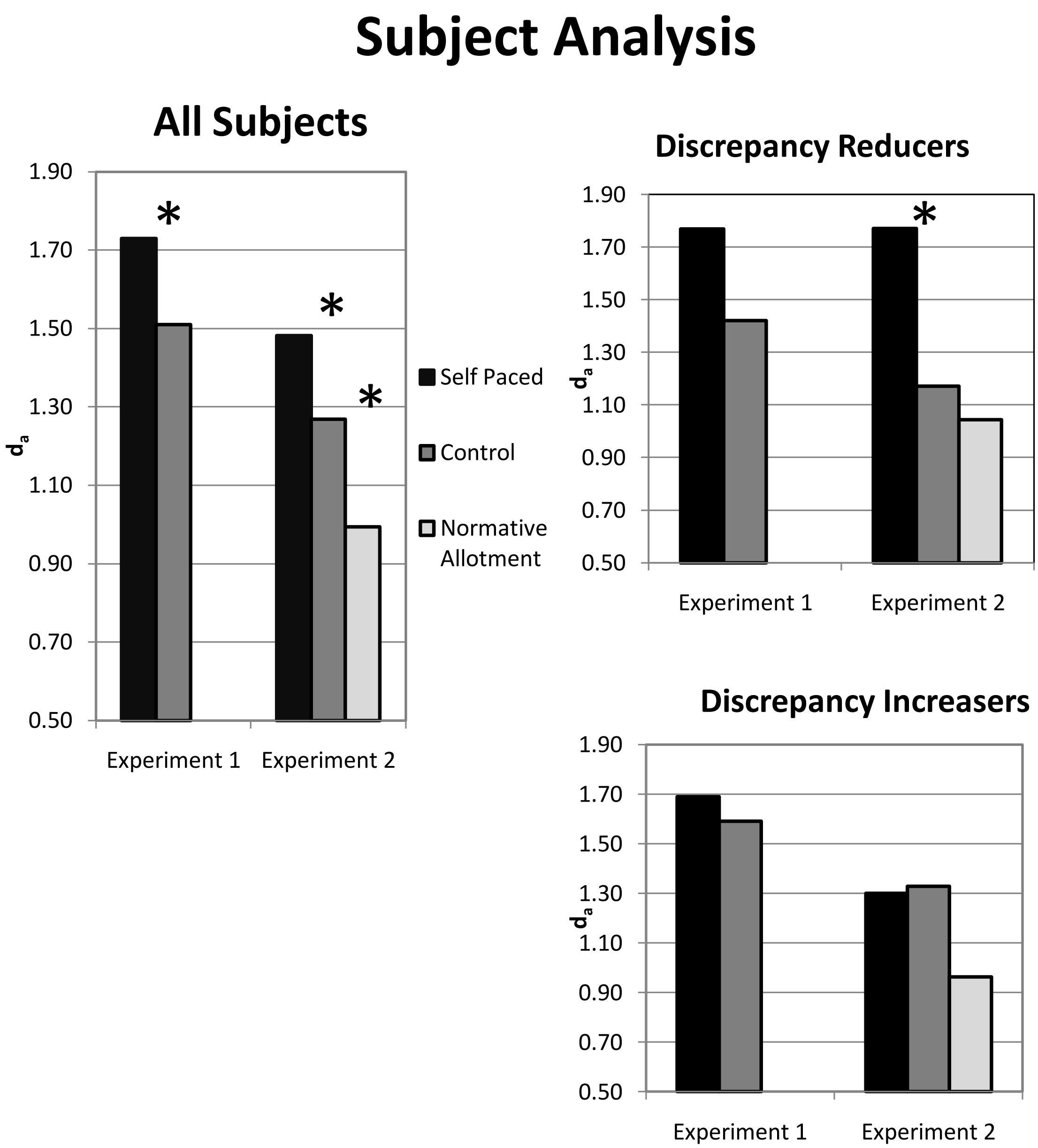


Figure 2. Left panel: Item discriminability as a function of condition based on an item analysis. Right panels: Discriminability for items as calculated by discrepancy reducers top) and discrepancy increasers (bottom).

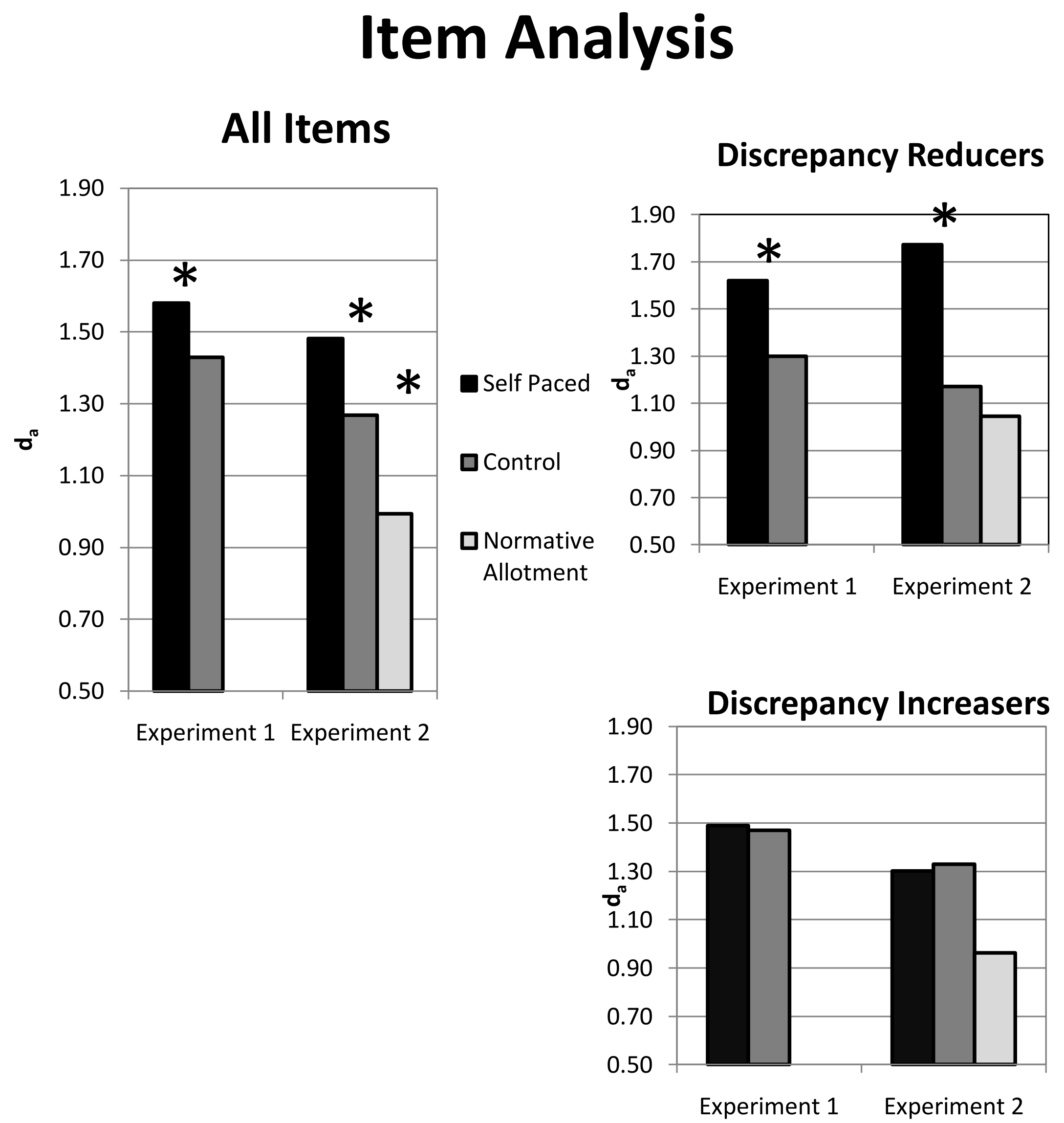
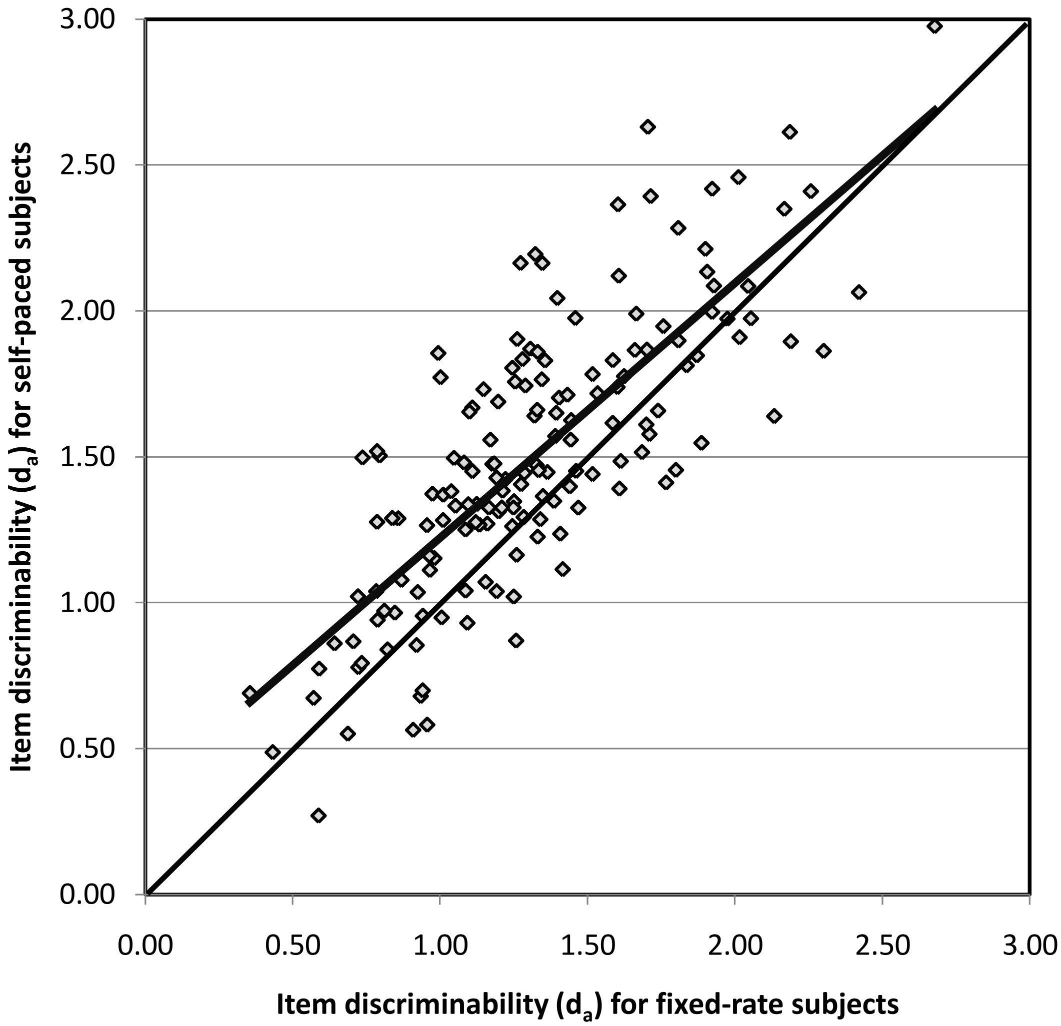


Figure 3. Item discriminability as computed from self-paced subjects plotted against item discriminability as computed from control (fixed-rate) subjects.



REFERENCES

Baker, L., & Anderson, R. I. (1982). Effects of inconsistent information on text processing: Evidence for comprehension monitoring. *Reading Research Quarterly*, *17*, 281-294.

Belmont, J. M., & Butterfield, E. C. (1971). Learning strategies as determinants of memory deficiencies. *Cognitive Psychology, 2,* 411–420.

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

Benjamin, A. S. (2005). Response speeding mediates the contribution of cue familiarity and target retrievability to metamnemonic judgments. *Psychonomic Bulletin & Review, 12*, 874-879.

Benjamin, A. S. (2008). Memory is more than just remembering: Strategic control of encoding, accessing memory, and making decisions. In A. S. Benjamin & B. H. Ross (Eds.), *The Psychology of Learning and Motivation: Skill and Strategy in Memory Use* (Vol. 48; pp.175-223). London: Academic Press.

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

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

Benjamin, A. S., Diaz, M. L., & Wee, S. (2009). Signal detection with criterion noise: Applications to recognition memory. *Psychological Review*, 116, 84-114.

Dunlosky, J., & Connor, L. T. (1997). Age differences in the allocation of study time account for age differences in memory performance. *Memory & Cognition*, *25*, 691-700.

Dufresne, A., & Kobasigawa, A. (1988). Developmental differences in children's allocation of study time. *The Journal of Genetic Psychology*, *149*, 87-92.

Dufresne, A., & Kobasigawa, A. (1989). Children's spontaneous allocation of study time: Differential and sufficient aspects. *Journal of Experimental Child Psychology*, *47*, 274–296.

Dunlosky, J., & Nelson, T. O. (1992). Importance of the kind of cue for judgements of learning (JOLs) and the delayed JOL effect. *Memory & Cognition*, *20*, 374-380.

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

Dunning, D., Johnson, K., Ehrlinger, J., & Kruger, J. (2003). Why people fail to recognize their own incompetence. *Current Directions in Psychological Science*, *12*, 83-87.

Finley, J. R., Tullis, J. G., & Benjamin, A. S. (2009). Metacognitive control of learning and remembering. In M. S. Khine & I. M. Saleh (Eds.) New Science of Learning: Cognition, Computers and Collaboration in Education. New York: Springer Science & Business Media.

Flavell, J. H. (1979). Metacognition and cognitive monitoring: A new area of cognitive-development theory. *American Psychologist*, *34*, 906-911.

Glanzer, M., & Bowles, N. (1976). Analysis of the word frequency effect in recognition memory. *Journal of Experimental Psychology: Human Learning & Memory*, *2*, 21-31.

Green, D., & Swets, J. (1966). *Signal detection theory and psychophysics*. New York: John Wiley and Sons.

Guttentag, R., & Carroll, D. (1998). Memorability judgments for high- and low-frequency words. *Memory & Cognition*, *26*, 951-958.

Hacker, D. J., Bol, L., Horgan, D., & Rakow, E. (2000). Test prediction and performance in a classroom context. *Journal of Educational Psychology*, *92*, 160-170.

Karpicke, J. D., & Roediger, H. L. (2008). The critical importance of retrieval for learning. *Science*, *319*, 966-968.

Kobasigawa, A., & Metcalf-Heggart, A. (1993). Spontaneous allocation of study-time by first- and third- grade children on a simple memory task. *The Journal of Genetic Psychology*, *154*, 223-235.

Koriat, A., Bjork, R. A., Sheffer, L., & Bar, S. K. (2004). Predicting one's own forgetting: The role of experience-based and theory-based processes. *Journal of Experimental Psychology: General*, *143*, 643-656.

Koriat, A., Ma'ayan, H., Nussinson, R, (2006). The intricate relationships between monitoring and control in metacognition: Lessons for the cause-and-effect relation between subjective experience and behavior. *Journal of Experimental Psychology: General*, 135, 36-60.

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

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

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

Macmillan, N. A., & Creelman, C. D. (2005). Detection Theory (2nd ed.). Mahwah, New Jersey: Lawrence Erlbaum.

Maki, R.H., & Serra, M. (1992). The basis of test predictions for text material. *The Journal of Experimental Psychology: Learning, Memory, and Cognition*, *18*, 116-126.

Mazzoni, G., & Cornoldi, C. (1993). Strategies in Study Time Allocation: Why is study time sometimes not effective? *Journal of Experimental Psychology: General*, *122*, 47-60.

Matzen, L. E. & Benjamin, A. S. (2009). Remembering words not presented in sentences: How study context changes patterns of false memories. *Memory & Cognition*, *37*, 52-64.

Metcalfe, J. (2009). Metacognitive judgments and control of study. *Current Directions in Psychological Science*, *18*, 159-163.

Metcalfe, J., & Finn, B. (2008). Evidence that judgments of learning are causally related to study choice. *Psychonomic Bulletin & Review*, *15*, 174-179.

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

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

Metcalfe, J., Schwartz, B. L., & Joaquim, S. G. (1993). The cue-familiarity heuristic in metacognition. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, 851-861.

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

Nelson, T. O., & Narens, L. (1994). Why investigate metacognition? In J. Metcalfe & A. P. Shimamura (Eds.), Metacognition: Knowing about knowing (pp. 195-206). Cambridge, MA: MIT Press.

Roediger, H. L., & Karpicke, J. D. (2006). Test-enhanced learning: Taking memory tests improves long term retention. *Psychological Science*, *17*, 249-255.

Shaughnessy, J. J. (1981). Memory monitoring accuracy and modification of rehearsal strategies. *Journal of Verbal Learning and Verbal Behavior*, *20*, 216-230.

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 Experiment Psychology: Learning, Memory, and Cognition*, *26*, 204-221.

Thiede, K. W., Anderson, M. C. M., & Therriault, D. (2003). Accuracy of metacognitive monitoring affects learning of texts. *Journal of Educational Psychology, 95*(1), 66-73.

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, and Cognition*, *35,* 1352-1358.

Wilson, M.D. (1988) The MRC Psycholinguistic Database: Machine Readable Dictionary, Version 2. *Behavioural Research Methods, Instruments and Computers, 20,* 6-11.

Zacks, R. T. (1969). Invariance of total learning time under different conditions of practice. *Journal of Experimental Psychology*, *32,* 441-447.

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.