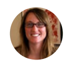
Agenda-Based Regulation of Study-Time Allocation: When Agendas Override Item-Based Monitoring

Article *in* Journal of Experimental Psychology General · September 2009

DOI: 10.1037/a0015928 · Source: PubMed

Robert Ariel
Virginia Wesleyan University



Heather Bailey
Kansas State University

Journal of Experimental Psychology: General
2009, Vol. 138, No. 3, 432–447

0096-3445/09/$12.00 DOI: 10.1037/a0015928

Agenda-Based Regulation of Study-Time Allocation: When Agendas Override Item-Based Monitoring

Robert Ariel, John Dunlosky, and Heather Bailey
Kent State University

Theories of self-regulated study assume that learners monitor item difficulty when making decisions about which items to select for study. To complement such theories, the authors propose an agenda-based regulation (ABR) model in which learners' study decisions are guided by an agenda that learners develop to prioritize items for study, given their goals and task constraints. Across 4 experiments, the authors orthogonally manipulated 1 task constraint—the reward structure of the task—with objective item difficulty, so that learners could use either item difficulty or potential reward in deciding how to allocate their study time. Learners studied items, were tested, and then selected half the items for restudy. As predicted by the ABR model, reward structure drove item selection more than did item difficulty, which demonstrates learners' agendas can override the effects of monitoring item difficulty in the allocation of study time.

Keywords: study-time allocation, metacognition, agenda-based regulation, planning, item difficulty

How do people allocate time while studying? And, is their allocation of time optimal? These questions have guided metacognitive research on self-regulated learning for over 3 decades (for a review, see Son & Kornell, 2008). Although researchers have not yet offered definitive answers to either question—especially the latter one—some consensus has arisen about how people allocate time while studying. In particular, previous research on self-regulated study has consistently demonstrated that learners use memory monitoring to control their study behavior (e.g., Son & Metcalfe, 2000; Thiede & Dunlosky, 1999). Monitoring is viewed as a driving force for study regulation because efficient regulation requires the ability to accurately monitor memory for study materials to make decisions about which materials one needs to learn (Dunlosky, Hertzog, Kennedy & Thiede, 2005; Thiede, Anderson, & Therriault, 2003). Monitoring is central to current theories of self-regulated study, which includes specific predictions about how learners will use memory monitoring to allocate study time (e.g, Dunlosky & Hertzog, 1998; Metcalfe & Finn, 2008; Nelson & Leonesio, 1988; Son & Metcalfe, 2000; Thiede & Dunlosky, 1999;). That is, these theories focus on the relation between monitoring and control of study.

In the present article, we investigate a complementary approach to understanding self-regulated study, which we call *agenda-based regulation* (ABR). According to this approach, learners use agendas to make decisions about how to allocate study time when faced with certain task constraints. At least in some cases, learners may violate the monitoring-control relations normally observed when learners' control of study is guided by monitoring. Before describing ABR in detail, we briefly review previous research and theory concerning people's allocation of study time.

Most theories and research on self-regulated study have focused primarily on the influence of item difficulty on the control of study time. A literature review conducted by Son and Metcalfe (2000) highlighted this emphasis on item difficulty. They presented the results of 31 experiments. Every experiment reported how learners regulate their study relative to perceived or objective item difficulty. Of the 46 experimental conditions in these experiments, 35 conditions supported the finding that learners allocate more time to studying difficult items, 3 conditions indicated that learners spend more time studying intermediate items, and 8 conditions indicated that learners did not show a preference for either difficult or easy items. Recent research has continued to highlight the importance of perceived item difficulty for study regulation, indicating that given various task constraints (e.g., amount of time available for study), learners will allocate more time either to items that are perceived as easy to learn or to items that are perceived as difficult to learn (Dunlosky & Thiede, 2004; Koriat, Ma'ayan, & Nussinson, 2006; Kornell & Metcalfe, 2006; Metcalfe & Kornell, 2003; Metcalfe & Kornell, 2005; Son & Sethi, 2006; Souchay & Isingrini, 2004; Thiede & Dunlosky, 1999).

Given such evidence, current theories of self-regulated study have been inspired by the monitoring-affects-control hypothesis proposed by Nelson and Leonesio (1988), in which it is assumed that people use monitoring of item difficulty to control study. Exactly how monitoring is used to regulate study, however, differs across specific theories. To illustrate this point, consider two theories that have received extensive scrutiny in the field. The discrepancy reduction theory, assumes that learners regulate their

Robert Ariel, John Dunlosky, and Heather Bailey, Psychology Department, Kent State University.

We thank the RADlab for many helpful discussions on this project, and Mary Pyc and Keith Thiede for commenting on drafts of this article. Special thanks go to Jennifer Wiley, who guided Robert Ariel's senior thesis on self-regulated learning at the University of Illinois at Chicago. Critiques of this article from Janet Metcalfe, Alan Castel, and Roddy Roediger were especially constructive and were appreciated.

Correspondence concerning this article should be addressed to John Dunlosky, Psychology Department, Kent State University, Kent, OH 44242. E-mail: jdunlosk@kent.edu

study to reduce the discrepancy between their perceived state of learning for a given item and their desired state of learning (Dunlosky & Hertzog, 1998). Accordingly, learners are expected to allocate more study time to items perceived as more difficult to learn than to items perceived as easier to learn because the former would initially be further from the norm of study and should require more study time to reach it. By contrast, the region of proximal learning (RPL) theory proposes that learners allocate their study to items that fall within their RPL (Metcalfe, 2002; Metcalfe & Kornell, 2005). This region refers to how readily a particular item can be learned. Some items may already be learned, whereas others may be too difficult to learn. The remaining items are within a learner's RPL, and the RPL theory predicts that a learner will initially allocate study time to these items. These theories make different predictions under some conditions. For instance, the discrepancy reduction theory always predicts learners will use the most time studying the more difficult items; by contrast, if the most difficult items are too difficult to learn, the RPL theory predicts that learners will initially use the most time studying easier items. Of most importance for now, however, is that these theories share one important feature: Item difficulty is a driving force behind decisions on how learners' allocate study time.

The model of ABR emphasizes the core role of agendas in study regulation. According to this model, learners develop an agenda on how to allocate time to various study items and use this agenda when selecting items for study. Like many other theories of regulation (e.g., Benjamin, 2007; Carver & Scheier, 2000; Pintrich, 2000), the ABR model assumes that study regulation is goal oriented. More specifically, our proposal is that learners attempt to allocate study to maximize the likelihood of obtaining task goals in the most efficient manner possible (Thiede & Dunlosky, 1999). Accordingly, an agenda is constructed to achieve this goal, given a variety of factors, including the task constraints and characteristics of the individual learner (cf. task and cognitive conditions, respectively, from Winne & Hadwin, 1998).

Developing an agenda involves prioritizing study materials, so that study time can be allocated to achieve task goals. The agenda that learners construct describes their decision criteria for item selection during study, and these decision criteria are chosen in an attempt to achieve task goals as efficiently as possible, given the current task constraints. Task constraints that affect how this agenda is constructed include the total time allotted for study, reward structure of the task, and task difficulty. The total amount of time for study impacts what study decisions are most efficient in a given context. For instance, when the amount of time for study is limited, it would not be efficient to study difficult material when acquisition of such material would require too much time. This particular task constraint would cause a learner to construct an agenda that involves giving easier study items high priority for learning (Dunlosky & Thiede, 2004). Reward structure of the task may also influence what constitutes efficient study decisions and may be particularly important in academic settings, such as when students realize that some materials will be worth more points on an exam or are more likely to be tested. Efficient regulation of study in this context would require developing an agenda that aims to maximize potential reward by allocating study to items that are worth more points or have a high likelihood of being tested. Certainly, factors other than reward structure may also influence agenda construction (for a review of other factors, such as learner characteristics, see Greene & Azevedo, 2007), but a key point here is that agendas may be constructed that will shift learners' attention from monitoring item difficulty to other factors (e.g., reward), as they allocate study time.

According to the ABR model, learners make decisions about whether to select an item for study by comparing each item with the decision criteria set forth by the learner's agenda. When a study item meets the criteria, it is selected for study, and when it does not, learners choose to restudy the item. Even with an agenda that specifies criteria for selection, the agenda may not always be executed. One limitation is that the ABR model assumes that learners' agenda execution is a top-down process that requires cognitive resources. In particular, an agenda will often include a goal and the particular criteria set for selecting items, and its execution may also involve keeping track of the dynamics of the ongoing learning task, such as which items have already been selected for restudy, how many items have (versus have not) been well learned, and so forth. Thus, ABR will often require learners to keep various kinds of information active in working memory as the agenda is executed. When cognitive resources are limited, top-down ABR may break down (Sobel, Gerrie, Poole, & Kane, 2007), such as when a task environment is distracting or when execution exceeds an individual's working-memory capacity (Barret, Tugade, & Engle, 2004). Thus, although the ABR model predicts that people's agendas can directly influence item selection, this influence may be diminished in some situations.

ABR has been included in the hierarchical model of self-regulated study proposed by Thiede and Dunlosky (1999). In this model, prestudy planning occurs at a superordinate level and influences which items learners will select for restudy, and then, control is transferred to a subordinate level at which self-paced study time is controlled by discrepancy reduction. The superordinate and subordinate levels of control are inextricably linked, and discrepancy reduction is viewed as foundational to study-time regulation. The contribution of discrepancy reduction to any kind of study regulation has been questioned (e.g., Metcalfe, 2002; Metcalfe & Kornell, 2005; but see Benjamin & Bird, 2005), which substantially limits the explanatory power of the hierarchical model. By contrast, the ABR model does not presuppose that discrepancy reduction drives allocation of study but instead emphasizes the potential role of agendas at all phases of regulation. Thus, given the limitations of the hierarchical model, we pursue ABR more broadly, with the intent to more precisely specify and explore the contributions of agendas to self-regulated study.

More generally, the idea that agendas influence memory performance is also not novel. In the source monitoring framework (Johnson, Hashtroudi, & Lindsay, 1993), agendas influence what information is activated during retrieval and how that information is evaluated (Mitchell et al., 2008). Also, younger and older adults appear to adopt different learning agendas under some conditions, which affects how they weigh the importance of information and their task performance (Castel, 2005; Castel, 2007; Castel, Benjamin, Craik, & Watkins et al., 2002; Castel, Farb, & Craik, 2007). More specifically, Castel et al. (2002) examined the effects of varying the point value for words on list recall. Words were presented in lists for study, and each word was assigned a point value ranging from 1 to 12 for correct recall. Participants were instructed to try to earn as many points as possible during the recall phase of the experiment. Younger and older adults recalled more words with higher point values (words ranging from 10 points to 12 points) than words with lower point values. That is, learners were using an agenda to strategically control their retrieval processes, so as to increase their final recall score (for similar findings, see Castel et al., 2007). Despite research highlighting the importance of agendas in the regulation of memory retrieval (see also, Benjamin, 2007; Reder, 1988; Schunn & Reder, 1998), ABR has not been systematically investigated in research on the regulation of study time.

Although an ABR model emphasizes the importance of agendas in regulating study decisions, agendas and monitoring jointly influence the allocation of study time. First, prior to applying the criteria for the agenda, learners presumably monitor their state of learning for a given item to evaluate whether they have already learned it, and if they have, they choose not to restudy it. This assumption—that learners will not study items that they judge they already know—has been supported by numerous studies (Dunlosky & Hertzog, 1997; Kornell & Metcalfe, 2006; Metcalfe & Kornell, 2005; Nelson, Dunlosky, Graf, & Narens, 1994). Assuming the current item has not been learned, learners consider their agenda in making decisions about whether to allocate more time to it. This agenda may also be informed by feedback from monitoring. For instance, as mentioned above, when learners have extreme time constraints, their agenda may include studying the easiest unknown items that can be quickly learned because sufficient time is not available to learn difficult items. Accordingly, learners would evaluate the ease of learning for each item and use this monitoring to allocate study time (cf. Son & Metcalfe, 2000; Thiede & Dunlosky, 1999). Even though monitoring may support ABR, however, ABR may not require or elicit monitoring memory for individual items, such as when the criteria for selection focus on aspects of the task that are not relevant to item difficulty. To illustrate this point, consider the following example, which is common in educational settings to which theories of self-regulated study are in part intended to generalize. Students are preparing for an examination, and although they know that some material will normatively be more difficult to learn than will others, they also realize that some material is more likely to appear on the upcoming test. Certain topics may have been emphasized more by the class instructor, and the instructor may have made it explicit that some items will likely be tested and that others will likely not be. In such cases, students are expected to develop an agenda that will prioritize for study the material that is most likely to be tested, regardless of the normative difficulty of learning the materials.

Overview of Experiments

The goal in the present experiments is to empirically evaluate the hypothesis that regulation of study time can be driven by the reward structure of a task, as predicted by the ABR model. In all four experiments, participants studied word pairs in which half the pairs were objectively easy and half were objectively difficult. Initial study was experimenter paced. Following initial study of the items, a preselection recall test was administered to assess which responses participants could recall prior to restudy. Finally, participants were allowed to select half the pairs (presented simultaneously in an array) for restudy.

Two separate manipulations were administered to the task materials that were hypothesized to affect the agenda that learners would construct to regulate study. The two manipulations were (a) the percentage likelihood that each study pair would be tested (Experiment 1) and (b) the number of points that would be awarded for correct recall of a given response (Experiments 2–4). Both manipulations reflect variations in task constraints that we expected would cause learners to develop different agendas for their study regulation. As discussed below, these manipulations were structured in the task to create tension between use of an agenda (i.e., selecting items that will yield the highest reward) and use of item difficulty to regulate study.

Experiment 1

In Experiment 1, participants studied word pairs that were presented as having a higher percentage likelihood of appearing on a final test (90%), for objectively easy items, and a lower percentage likelihood (30%), for objectively difficult items (high-likelihood easy group); a higher percentage likelihood (90%) for difficult items and a lower percentage likelihood (30%) for easy items (high-likelihood difficult group); or an equal percentage likelihood (30%) for both easy and difficult items (constant-likelihood group). We chose these values because we expected that the difference in percentage likelihood of appearing on the test (60% difference) would be noticed by participants but would not be so extreme (100% versus 0%) as to produce a trivial outcome. Also, we wanted to use one of the two percentage likelihood values from the experimental groups for items in the constant-likelihood group (e.g., instead of using a 50% constant likelihood); our choice of 30% (versus 90%) was admittedly arbitrary because we did not believe a constant value would be a factor in the selection of items. Note, however, that the most critical comparisons occur between the two high-likelihood groups.

Participants were expected to develop an agenda in an attempt to maximize performance. Thus, in the ABR model, it is predicted that during item selection, participants will allocate their study to the unrecalled items that have the higher percentage likelihood of appearing on a final test, regardless of whether those items are objectively difficult or easy to learn. If this agenda is driving item selection as expected, a crossover interaction should emerge, in which participants in the high-likelihood easy group are more likely to select easier (vs. more difficult) items for restudy, whereas those in the high-likelihood difficult group are more likely to select the more difficult items.

Without an auxiliary assumption (which is considered in the introduction to Experiment 4), the two other theories of self-regulated study do not predict this crossover interaction. In particular, the discrepancy reduction theory predicts that learners will always select the most difficult (unlearned) items for study, and the RPL theory predicts that learners will select unlearned items in their RPL. Because item difficulty and RPL should be identical across all groups, each theory then predicts that participants will allocate restudy time identically across the likelihood groups. Of course, these two theories may predict different patterns of allocation, assuming that participants' RPL does not consist of the most difficult items. The important point here, however, is that both theories predict that study allocation should be the same across the three groups, whether it be allocated to the most difficult items (discrepancy reduction theory) or to the RPL.

Method

Participants. Sixty participants from Kent State University participated for course credit in Introductory Psychology. Participants were randomly assigned to either the high-likelihood easy group ($n = 20$), the high-likelihood difficult group ($n = 20$), or the constant-likelihood control group ($n = 20$).

Materials. Thirty noun–noun paired associates were used. The difficulty of the items was manipulated within-groups, so that 15 paired associates were easy and 15 were difficult. The easy items consisted of concrete–concrete pairs (e.g., *book–hammer*), and the difficult items consisted of abstract–abstract pairs (e.g., *liberty–velocity*). The experiment was administered with E-prime software (Schneider, Eschman, & Zuccolotto, 2001).

Procedure. Participants completed the experiment individually (each participant at his or her own computer) in a group computer lab. All instructions were displayed on the computer screens; all participants were instructed that they would be studying for an upcoming test. Participants in the constant-likelihood group were told that the likelihood any given item would be present on the final test was 30%. Participants in the high-likelihood easy and high-likelihood difficult groups were instructed that the likelihood any given item would be present on the final test would be 30% or 90%. In the high-likelihood easy group, easy items were presented during study phases as 90% likely to appear on the final test, and difficult items were presented as 30% likely to appear on the final test. In the high-likelihood difficult group, difficult items were presented as 90% likely to appear on the final test, and easy items were presented as 30% likely to appear on the final test. Participants were not informed that these percentages were assigned according to item difficulty.

During the initial study phase of the experiment, each paired associate was presented individually for 6 s. The percentage likelihood of being tested on each item (either 30% or 90%) appeared above each pair. Immediately after studying a pair, participants made an ease-of-learning (EOL) judgment; the stimulus of a pair alone was presented along with this prompt: "How difficult do you feel it would be to learn this pair at a later time? 1 (meaning *very easy*), 2, 3, 4, 5, or 6 (meaning *very difficult*)." EOL judgments were made by participants viewing only the stimulus of each pair. Participants typed their response, and then, the next item was presented for study. Given that EOL judgments were not critical for evaluating our central predictions and were consistent with objective item difficulty, we do not discuss them further. Interested readers can obtain these data by contacting John Dunlosky. After studying all of the items, a paired-associate recall test was administered. In the test, participants were presented with the stimulus of each word pair individually and were prompted to type the corresponding target.

Following testing, participants were allowed to choose 15 items for restudy. Participants selected items from a 5 × 6 array. In each of the 30 cells of the array, the percentage likelihood of the pair appearing on the final test and the first word of that pair were presented (e.g., 90% liberty; 30% book). In the top left corner of each cell, either a letter from A to Z or a number from 1 to 4 was presented for reference. Participants selected items for restudy by typing the letter or number that corresponded to the pair they wanted to restudy. When an item was selected for restudy, it was presented immediately, and participants were allowed to study until they were ready to select the next pair for restudy. After choosing 15 items for restudy, a final paired-associate recall test was administered; all items were tested, regardless of the initial manipulation concerning the likelihood of appearing on the test.

Results

We begin our analyses by examining recall performance prior to the selection phase (i.e., preselection recall), to demonstrate that objective item difficulty was successfully manipulated. Most important for evaluating the ABR model, we then report analysis of item selection for items that were not recalled during preselection recall. For completeness, the same analyses were then repeated for items that were correctly recalled during the preselection recall phases. Finally, we present self-paced study times for unrecalled items and then report analysis of final recall.

Preselection recall. Proportion of correct recall on the preselection recall trials was computed for each participant, and a mean was then computed across participants' values (Table 1). A 2 (item difficulty: easy vs. difficult) × 3 (likelihood group: high-likelihood easy, high-likelihood difficult, and constant likelihood) analysis of variance (ANOVA) revealed a significant main effect for item difficulty, which confirms that participants in each group correctly recalled more easy items than difficult items, $F(1, 57) =$

Table 1
Proportion of Items Correctly Recalled During Preselection Recall

	Item difficulty			
	Easy items		Difficult items	
Group	*M*	*SE*	*M*	*SE*
Experiment 1				
High-likelihood easy	.57	.06	.24	.05
High-likelihood difficult	.37	.06	.10	.03
Constant-likelihood	.50	.05	.16	.04
Experiment 2				
High-reward easy	.44	.05	.12	.02
High-reward difficult	.61	.05	.17	.04
Constant-reward	.47	.05	.20	.04
Experiment 3				
High-reward easy	.46	.14	.13	.03
High-reward difficult	.53	.05	.20	.04
Constant-reward	.55	.04	.19	.05
Experiment 4				
Sequential format				
High-reward easy	.46	.06	.15	.02
High-reward difficult	.41	.05	.17	.04
Simultaneous format				
High-reward easy	.42	.07	.13	.03
High-reward difficult	.41	.05	.17	.03

Note. Values are means across an individual participant's proportion of correct recall.

145.35, $MSE = 0.02$, $p < .001$, $\eta_p^2 = .72$. The main effect for group was significant, $F(2, 57) = 3.69$, $MSE = 0.08$, $p < .05$, $\eta_p^2 = .12$, and a Tukey honestly significant difference (HSD) post hoc test indicated that more responses were correctly recalled by the high-likelihood easy group than by the high-likelihood difficult group. This effect was unexpected, and most important, its presence does not influence the a priori predictions of the theories. Moreover, this unexpected group effect did not arise in the other experiments, so we do not consider it further. The interaction was not significant, $F(2, 57) = 0.61$, $MSE = 0.01$, $p = .55$.

Item selection for unrecalled items. As is most critical for evaluating predictions from the theories, we examined item selection for those items that were not recalled during preselection recall because current theories indicate that participants will favor these items for restudy (e.g., Metcalfe & Kornell, 2005). For items that were not correctly recalled during preselection recall, we computed the mean proportion selected for restudy by each participant. Means across participants' values are presented in Figure 1. As evident from inspection of this figure, the predictions of the ABR model were confirmed. Participants in the high-likelihood easy group selected more easy items for restudy than difficult items, and participants in the high-likelihood difficult group selected more difficult items than easy items. Consistent with these observations, a 2 (item difficulty) × 3 (group) ANOVA revealed a main effect for group, $F(2, 57) = 4.51$, $MSE = 0.07$, $p < .05$, $\eta_p^2 = .14$, which was qualified by a reliable interaction between item difficulty and group, $F(2, 57) = 12.13$, $MSE = 1.44$, $p < .001$, $\eta_p^2 = .30$. The effect of item difficulty was not significant, $F(1, 57) = 0.61$, $MSE = 0.12$, $p = .44$.

To unpack the interaction and to evaluate the predictions from the various theories, we conducted a series of follow-up ANOVAs. First, a 2 (item difficulty) × 2 (high-likelihood easy group vs. high-likelihood difficult group) ANOVA was conducted to compare the high-likelihood easy group and the high-likelihood difficult group. As predicted, a significant crossover interaction between item difficulty and group was obtained, $F(1, 38) = 26.67$, $MSE = 2.8$, $p < .001$, $\eta_p^2 = .41$. Next, a 2 (item difficulty) × 2 (high-likelihood easy group vs. constant-likelihood group) ANOVA was conducted to compare the high-likelihood easy group and constant-likelihood group. The interaction between item difficulty and group was not significant for these groups, $F(1, 38) = 2.72$, $MSE = 0.36$, $p = .11$. Finally, the high-likelihood difficult group was compared with the constant-likelihood group, which revealed a significant Item Difficulty × Group interaction, $F(1, 38) = 9.75$, $MSE = 1.15$, $p < .01$, $\eta_p^2 = .20$.

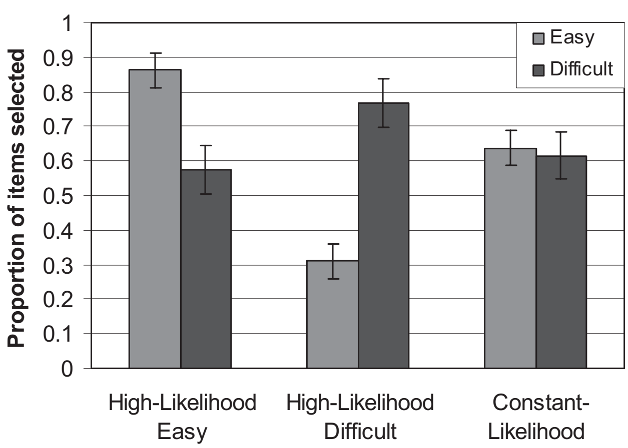


Figure 1. Mean proportion of items selected for restudy (Experiment 1) for items that were not correctly recalled during preselection recall. Error bars represent standard error of the mean.

Item selection for recalled items. The mean proportion of recalled items selected for restudy is presented in Table 2. A 2 (item difficulty) × 3 (group) ANOVA revealed no main effects for item difficulty, $F(1, 55) = 0.09$, $MSE = 0.01$, $p = .77$, or for group, $F(2, 55) = 0.02$, $MSE = 0.12$, $p = .98$. The interaction approached significance, $F(2, 55) = 2.81$, $MSE = 0.44$, $p = .07$.

Self-paced study for unrecalled items. Mean self-paced times for items that were not recalled during preselection recall were computed, to examine whether reward influenced the length of time participants spent studying items they selected for restudy. As evident from inspection of Table 3, participants persisted longer in studying items that had a high-likelihood of appearing on a final test. Consistent with this observation, a 2 (item difficulty) × 3 (likelihood group) ANOVA yielded no effects for item difficulty, $F(1, 56) = 0.49$, $MSE = 8.55$, $p = .49$, or for reward group, $F(2, 56) = 0.29$, $MSE = 8.11$, $p = .75$, but the interaction was significant, $F(2, 56) = 4.55$, $MSE = 38.91$, $p < .05$, $\eta_p^2 = .14$.

Follow-up analyses were conducted to unpack this interaction. A 2 (item difficulty) × 2 (likelihood group) ANOVA comparing the high-likelihood easy and difficult groups revealed a significant interaction, $F(1, 38) = 6.59$, $MSE = 77.47$, $p < .01$, $\eta_p^2 = .15$. As with item selection (Figure 1), this crossover interaction revealed that participants spent more time studying easy items when they had a high-likelihood of being tested, whereas participants spent more time studying difficult items when those items had a high-likelihood of being tested. Comparisons of the high-likelihood easy group and the constant-likelihood group did not reveal an interaction, $F(1, 37) = 1.68$, $MSE = 14.70$, $p = .20$. By contrast, comparisons of the high-likelihood difficult group and the constant-reward group revealed a significant interaction, $F(1, 37) = 4.66$, $MSE = 23.57$, $p < .05$, $\eta_p^2 = .11$, which indicated that participants tended to study difficult items longer when they had a high-likelihood of appearing on the test.

Final recall. Although less relevant to our current aims, we also examined the degree to which restudy boosted performance from the preselection recall phase to the test of final recall. Accordingly, we analyzed the increase in the proportion of items recalled from preselection recall to final recall (Table 4). (Note that all items appeared on the final test.)

A 2 (item difficulty) × 3 (likelihood group) ANOVA revealed no main effect for item difficulty, $F(1, 57) = 1.53$, $MSE = 0.02$, $p = .22$, or for likelihood group, $F(2, 57) = 0.04$, $MSE = 0.02$, $p = .96$. The Item Difficulty × Group interaction was significant, $F(2, 57) = 8.58$, $MSE = 0.19$, $p < .001$, $\eta_p^2 = .23$. To unpack this interaction, we first compared the two high-likelihood groups, using a 2 (item difficulty) × 2 (likelihood: easy vs. difficult) ANOVA. The interaction was significant, $F(1, 38) = 12.90$, $MSE = 0.35$, $p < .001$, $\eta_p^2 = .25$, which indicated that participants benefited from allocating more study time to the 90% likelihood items in the high-likelihood easy group and the high-likelihood difficult group.

Table 2
Proportion of Recalled Items Selected for Restudy

Group	Easy items *M*	Easy items *SE*	Difficult items *M*	Difficult items *SE*
Experiment 1				
High-likelihood easy	.36	.08	.33	.08
High-likelihood difficult	.24	.08	.50	.11
Constant-likelihood	.44	.08	.28	.08
Experiment 2				
High-reward easy	.38	.08	.24	.08
High-reward difficult	.14	.04	.38	.09
Constant-reward	.35	.08	.34	.08
Experiment 3				
High-reward easy	.53	.07	.28	.08
High-reward difficult	.20	.05	.46	.08
Constant-reward	.33	.05	.35	.08
Experiment 4				
Sequential				
High-reward easy	.32	.07	.31	.07
High-reward difficult	.25	.06	.39	.09
Simultaneous				
High-reward easy	.52	.09	.12	.05
High-reward difficult	.10	.03	.56	.09

Note. Values are the mean proportion of unrecalled items that were selected for restudy.

We also compared each of the two high-likelihood groups to the constant-likelihood group. Comparisons of the high-likelihood easy group and the constant-likelihood group did not reveal an interaction, $F(1, 38) = 1.11$, $MSE = 0.02$, $p = .30$. In contrast, comparisons of the high-likelihood difficult group and the constant-likelihood group revealed a significant interaction, $F(1, 38) = 10.41$, $MSE = 0.19$, $p < .01$, $\eta_p^2 = .22$. Overall, these outcomes indicate that learners' allocation of study time yielded labor-and-gain effects for (at least) difficult items slated for a higher likelihood of appearing on the test.

Table 3
Self-Paced Study Time for Unrecalled Items

Group	Easy items *M*	Easy items *SE*	Difficult items *M*	Difficult items *SE*
Experiment 1				
High-likelihood easy	5.76	1.21	4.25	0.72
High-likelihood difficult	3.74	1.12	6.16	0.98
Constant-likelihood	4.07	0.66	4.30	0.91
Experiment 2				
High-reward easy	6.48	0.82	4.67	0.64
High-reward difficult	4.44	1.12	7.67	1.03
Constant-reward	5.73	1.23	6.47	1.38
Experiment 3				
High-reward easy	3.50	0.87	3.52	0.61
High-reward difficult	2.63	0.76	6.13	0.14
Constant-reward	3.18	0.56	4.19	0.62

Note. Values are mean self-paced study-time for unrecalled items in seconds. Self-paced study times were not collected in Experiment 4.

Table 4
Increase in Proportion of Items Correctly Recalled From Preselection Recall to Final Recall

Group	Easy items *M*	Easy items *SE*	Difficult items *M*	Difficult items *SE*
Experiment 1				
High-likelihood easy	.21	.04	.14	.03
High-likelihood difficult	.07	.02	.26	.04
Constant-likelihood	.18	.03	.17	.03
Experiment 2				
High-reward easy	.26	.04	.13	.03
High-reward difficult	.07	.03	.32	.04
Constant-reward	.17	.03	.15	.04

Note. Values are the mean increase across an individual participant's proportion of correct recall. Final recall performance was not collected in Experiments 3 and 4.

Discussion

In accord with the predictions of the ABR model, participants in the high-likelihood groups were more likely to select items for restudy that had a high percentage likelihood of appearing on an upcoming test than to select items that had a low percentage likelihood. The resulting crossover interaction (Figure 1) poses problems for models of self-regulated study that emphasize item difficulty as driving the regulation of study (e.g., Dunlosky & Hertzog, 1998; Metcalfe & Kornell, 2005). For those items that participants selected for restudy, reward also influenced self-paced study time (cf. Dunlosky & Thiede, 1998): Participants allocated more time to the items that were more likely to appear on the test (Table 3). This effect suggests that agendas may influence self-paced study, and this effect cannot be readily explained by the hierarchical model of self-regulated study (Thiede & Dunlosky, 1999). Most important, although current theories of self-regulated study (based on item difficulty) may account for how people regulate study under some circumstances, none provides a complete explanation of the extant data.

A similar trend appeared in the patterns of item selection for items that were recalled prior to restudy (Table 2), as compared with those not recalled prior to restudy (Figure 1). That participants even selected for restudy previously recalled items was unexpected, given that people often choose not to restudy items they have already recalled (e.g., Dunlosky & Hertzog, 1997; Kornell & Bjork, 2008; Kornell & Metcalfe, 2006; Metcalfe & Kornell, 2005; Nelson, Dunlosky, Graf, & Narens, 1994). Even so, we recommend caution in interpreting the current pattern of selection for previously recalled items as inconsistent with previous research. For instance, preselection recall here occurred well prior to selection itself, so participants may have forgotten some of the previously recalled items before the selection phase, and hence, the pattern of selection effects for previously recalled items (Table 2) may reflect selection on a subset of items that had since been forgotten. Given the ambiguous nature of interpreting item selection for recalled items, we do not discuss this issue further in this article, but the data are presented in subsequent sections for completeness.

Experiment 2

In Experiment 2, our goal was to influence learners' agendas for study by manipulating a different aspect of the reward structure of the task—the number of points (1 point or 5 points) participants would receive for correct recall of items. This manipulation was structured the same as the percentage likelihood manipulation in Experiment 1, so as to create tension between use of potential reward and use of item difficulty in selecting items for restudy. According to the ABR model, participants will select unrecalled items with a high reward, regardless of whether this reward is slated with the objectively easier items or the more difficult ones.

Method

Participants. Sixty participants from Kent State University participated in this experiment for course credit in Introductory Psychology. Participants were randomly assigned to either the high-reward easy group ($n = 20$), the high-reward difficult group ($n = 20$), or the constant-reward group ($n = 20$).

Materials and procedure. The same materials and procedure used in Experiment 1 were used in Experiment 2, with the exception that probabilities were replaced with the amount of points a participant would receive for correctly recalling an item on the final test. The reward for learning items was manipulated across groups for study and restudy phases of the experiment. All paired associates in the constant-reward group were presented as being worth 1 point. For the high-reward easy group, easy items were worth 5 points and difficult items were worth 1 point. For the high-reward difficult group, the difficult items were worth 5 points and the easy items were worth 1 point. Once again, participants were not told that the points varied as a function of item difficulty. Prior to beginning the study phase, participants were instructed that their goal for the task was to try to earn as many points as possible during final recall.

Results

Preselection recall. Preselection recall was analyzed as in Experiment 1 (Table 1). A 2 (item difficulty) × 3 (reward group: high-reward easy, high-reward difficult, and constant reward) ANOVA revealed a significant main effect for item difficulty, $F(1, 57) = 284.87$, $MSE = 0.01$, $p < .001$, $\eta_p^2 = .83$, indicating that participants recalled more easy items than difficult items. The main effect for reward group was not significant, $F(2, 57) = 1.86$, $MSE = 0.06$, $p = .17$, although the interaction was significant, $F(2, 57) = 5.52$, $MSE = 0.07$, $p < .01$, $\eta_p^2 = .16$. The lack of a main effect of point reward is inconsistent with previous literature demonstrating that increasing incentives (operationalized as monetary reward) presented during an experimenter-paced study trial can boost subsequent recall performance (e.g., Eysenck & Eysenck, 1982; Loftus & Wickens, 1970; Weiner, 1966). One explanation for these effects of incentive is that people use more elaborative encoding for items slated for more reward, which in turn boosts performance (Eysenck & Eysenck, 1982). Although speculative, the lack of a reward effect in our experiment may be due to the fact that participants had enough time to benefit from elaboration (especially for easy items), regardless of the point reward. This possibility could be evaluated in future research, but given that our focus is on item selection, we do not discuss preselection recall further.

Item selection for unrecalled items. Item selection was computed as in Experiment 1, and as shown in Figure 2, reward drove item selection. The main effects for item difficulty, $F(1, 57) = 0.09$, $MSE = 0.08$, $p = .76$, and for reward group, $F(2, 57) = 0.13$, $MSE = 0.09$, $p = .88$, were not significant, but their interaction was significant, $F(2, 57) = 20.97$, $MSE = 1.66$, $p < .001$, $\eta_p^2 = .42$. As in Experiment 1, planned comparisons were conducted to unpack this interaction. First, the high-reward easy and high-reward difficult groups were compared. A 2 (item difficulty) × 2 (reward group) ANOVA revealed a significant crossover interaction between item difficulty and reward group, $F(1, 37) = 33.77$, $MSE = 2.83$, $p < .001$, $\eta_p^2 = .48$, which indicates that participants selected more of the high-reward items for study regardless of item difficulty. Next, a 2 (item difficulty) × 2 (reward group) ANOVA comparing the high-reward easy group with the constant-reward group also revealed a significant interaction, $F(1, 38) = 4.25$, $MSE = 0.29$, $p < .05$, $\eta_p^2 = .10$. Finally, the high-reward difficult group was compared with the constant-reward group, and the interaction between item difficulty and reward group was again significant, $F(1, 37) = 16.82$, $MSE = 1.32$, $p < .001$, $\eta_p^2 = .31$.

Item selection for recalled items. The mean proportion of recalled items selected for restudy is presented in Table 2. A 2 (item difficulty) × 3 (reward group) ANOVA indicated that the effect for item difficulty, $F(1, 57) = 0.35$, $MSE = 0.10$, $p = .56$, and for reward group, $F(2, 57) = 0.70$, $MSE = 0.14$, $p = .56$, were not significant. An interaction between item difficulty and group was significant, $F(2, 57) = 3.88$, $MSE = 0.38$, $p < .05$, $\eta_p^2 = .12$.

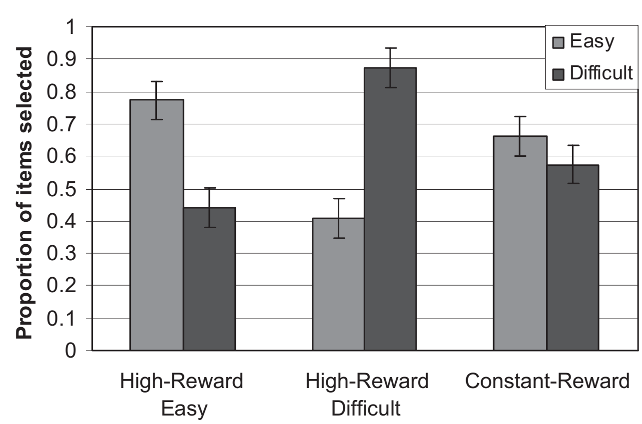


Figure 2. Mean proportion of items selected for restudy (Experiment 2) for items that were not correctly recalled during preselection recall. Error bars represent standard error of the mean.

Planned comparisons were conducted to unpack this interaction. Most important, the high-reward easy group was compared to the high-reward difficult group. A 2 (item difficulty) × 2 (reward group) ANOVA revealed a significant crossover interaction, $F(1, 38) = 8.10$, $MSE = 0.73$, $p < .01$, $\eta_p^2 = .18$, which paralleled the key outcome for selection of unrecalled items. Next, the high-reward easy group was compared with the constant-reward group, and no interaction was found, $F(1, 38) = 0.80$, $MSE = 0.09$, $p = .38$. Finally, the high-reward difficult group was compared with the constant-reward group, and the interaction approached significance, $F(1, 38) = 3.35$, $MSE = 0.32$, $p = .08$.

Self-paced study for unrecalled items. Mean self-paced study times for easy and difficult items are presented in Table 3. The effect for item difficulty approached significance, $F(1, 56) = 3.04$, $MSE = 4.97$, $p = .09$, and the effect for reward group was not significant, $F(2, 56) = 0.09$, $MSE = 3.33$, $p = .92$. The Item Difficulty × Reward Group interaction was significant, $F(2, 56) = 12.81$, $MSE = 63.63$, $p < .001$, $\eta_p^2 = .31$. To explore this interaction, we conducted the following planned comparisons. A 2 (item difficulty) × 2 (reward group) ANOVA comparing the high-reward groups revealed a significant crossover interaction, $F(1, 38) = 22.35$, $MSE = 127.26$, $p < .001$, $\eta_p^2 = .37$, which highlights that participants studied items that were slated for a high reward longer than items slated for a low reward. Comparisons of the constant-reward group with both the high-reward easy group, $F(1, 37) = 7.30$, $MSE = 31.75$, $p < .01$, $\eta_p^2 = .17$, and the high-reward difficult group, $F(1, 38) = 6.25$, $MSE = 30.25$, $p < .05$, $\eta_p^2 = .14$, yielded significant interactions.

Final recall. The increase in the proportion of items recalled on the preselection recall phase to final recall is presented in Table 4. The effects for item difficulty, $F(1, 57) = 1.10$, $MSE = 0.03$, $p = .30$, and reward group, $F(2, 57) = 0.64$, $MSE = 0.01$, $p = .53$, were not significant. The interaction was significant, $F(2, 57) = 14.62$, $MSE = 0.39$, $p < .001$, $\eta_p^2 = .34$, so we conducted follow-up ANOVAs. A 2 (item difficulty) × 2 (reward group) ANOVA comparing the high-reward groups revealed a significant interaction, $F(1, 38) = 24.86$, $MSE = 0.72$, $p < .001$, $\eta_p^2 = .40$. As in Experiment 1, this interaction demonstrates a labor-and-gain effect because participants benefited from allocating more study time to 5-point items than to 1-point items. Finally, we compared the constant-reward group with the two high-reward groups. The comparison between the constant-reward group and the high-reward easy group did not reveal a significant interaction, $F(1, 38) = 2.08$, $MSE = 0.05$, $p = .16$. However, the comparison involving the high-reward difficult group did reveal a significant interaction, $F(1, 38) = 15.64$, $MSE = 0.38$, $p < .001$, $\eta_p^2 = .30$. Given that final recall is not focal for evaluating the core predictions from the ABR model, we did not collect it in the next experiments, and we do not consider these analyses further until the General Discussion.

Discussion

The crossover interaction between the two high-reward groups (high-reward easy and high-reward difficult) and item difficulty (Figure 2) indicates that item difficulty alone is not driving regulation. These results provide further evidence that manipulating the reward structure of the task influences item selection during study, regardless of item difficulty. According to the ABR model, learners believe that selecting high-reward items (5-point items) will facilitate earning as many points as possible, so learners develop an agenda to prioritize high-reward items for restudy. Although intuitive and simple in nature, this agenda evidently overrides learners' typical response to focus restudy on either the most difficult items or the items in the RPL.

Experiment 3

In Experiment 3, we addressed some of the shortcomings of the previous experiments. In particular, participants selected items for study while viewing the stimulus of each pair and either the percentage likelihood it would appear on an upcoming test (Experiment 1) or the reward value for that item (Experiment 2). It is possible that participants were unlikely to monitor item difficulty during selection because variations in item difficulty were not as salient as variations in reward structure. Accordingly, during item selection in Experiment 3, participants were also shown the objective difficulty of each item. That is, in addition to the stimulus of each item and the reward value, either the word *easy* or the word *difficult* was presented.

Moreover, previous research has demonstrated that item relatedness consistently has a large (and similar) effect both on recall and on the magnitude of metamemory judgments (e.g., Carroll, Nelson, & Kirwan, 1997; Connor, Dunlosky, & Hertzog, 1997; Dunlosky & Matvey, 2001; Hertzog, Kidder, Powell-Moman, & Dunlosky, 2002; Koriat, 1997; Rabinowitz, Ackerman, Craik, & Hinchley, 1982). Thus, in Experiment 3, difficulty was manipulated with related word pairs (e.g., paint–artist) and unrelated word pairs (e.g., animal–library). Participants were instructed that easy items were easier to learn because the stimulus and response of the pair were conceptually related and that words in difficult pairs were conceptually unrelated. The goal of all these changes was to make item difficulty more salient during selection than in the first two experiments. If reward structure again overrides item difficulty in the allocation of study time, it cannot be attributed to participants being largely unaware of the variation in item difficulty across pairs.

Method

Participants. Seventy-five participants from Kent State University participated for course credit in Introductory Psychology ($n = 56$) or for U.S.$10 ($n = 19$). Participants were randomly assigned to either the high-reward easy group (total $n = 25$; paid $n = 6$), the high-reward difficult group ($n = 25$; paid $n = 9$), or the constant-reward group ($n = 25$; paid $n = 4$). Analyses conducted with (versus without) the paid participants yielded identical conclusions, so we included them in all analyses.

Materials. Thirty paired associates were used, and item difficulty was manipulated within-groups, so that 15 paired associates were easy and 15 were difficult. Objectively easy items were conceptually related word pairs (e.g., *paint–artist*), and objectively difficult pairs were unrelated word pairs (e.g., *animal–library*).

Procedure. The procedure was the same as the one used in Experiment 2, with the following modifications. First, participants were instructed that related word pairs would be easier to learn than would unrelated words pairs. Second, during the initial study phase, each word pair was presented for only 1.5 s. Similarly, after participants completed the initial study phase, they performed a 15 min distracter task—completing math problems. The goal in using a faster presentation rate and a distracter task was to ensure that preselection recall for related pairs would not be on the performance ceiling. Third, during the selection phase, item difficulty (easy or difficult) was also presented (along with the stimulus of a pair and the reward value) beside each pair. Instructions were presented next to the array that reminded participants that the easy items were related pairs and difficult items were unrelated pairs. Finally, given that final recall was not focal in our current evaluations of the ABR model, we did not collect final recall data in the final two experiments.

Results

Preselection recall. Mean recall is presented in Table 1. Participants recalled more easy items than difficult ones, $F(1, 71) = 184.76$, $MSE = 0.02$, $p < .001$, $\eta_p^2 = .72$. The effect for reward group, $F(2, 71) = 1.18$, $MSE = 0.08$, $p = .31$, and the Item Difficulty × Reward interaction, $F(2, 71) = 0.12$, $MSE = 0.01$, $p = .89$, were not significant.

Item selection for unrecalled items. Mean proportion of unrecalled items that participants selected for study was computed as in the previous experiments. As evident from inspecting Figure 3, participants selected more 5-point items than 1-point items, regardless of item difficulty. The main effect for item difficulty approached significance, $F(1, 72) = 3.37$, $MSE = 0.07$, $p = .07$, and the effect of reward group was not significant, $F(2, 72) = 1.31$, $MSE = 0.07$, $p = .28$. Most important, the interaction between these factors was significant, $F(2, 72) = 19.89$, $MSE = 1.44$, $p < .001$, $\eta_p^2 = .36$. As in previous experiments, planned comparisons were conducted to further examine this interaction.

A 2 (item difficulty) × 2 (reward group) ANOVA was conducted to compare the high-reward easy and high-reward difficult groups. The crossover interaction evident in Figure 3 was significant, $F(1, 48) = 39.71$, $MSE = 2.83$, $p < .001$, $\eta_p^2 = .45$. Next, the high-reward easy group was compared with the constant-reward group, which also revealed a significant interaction between item difficulty and reward group, $F(1, 48) = 5.93$, $MSE = 0.45$, $p < .05$, $\eta_p^2 = .11$. Finally, a 2 × 2 ANOVA was conducted to compare the high-reward difficult group and the constant-reward group, and again, the interaction was significant, $F(1, 48) = 14.73$, $MSE = 1.03$, $p < .001$, $\eta_p^2 = .24$.

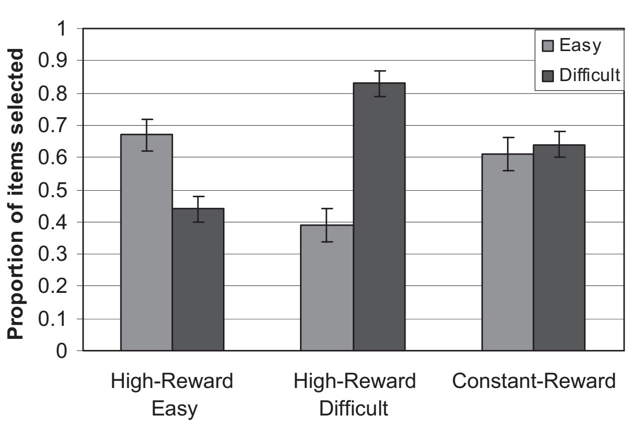


Figure 3. Mean proportion of items selected for restudy (Experiment 3) for items that were not correctly recalled during preselection recall. Error bars represent standard error of the mean.

Item selection for recalled items. Mean proportion of recalled items selected for restudy are presented in Table 2. A 2 (item difficulty) × 3 (reward group) ANOVA revealed no effect of item difficulty, $F(1, 71) = 0.04$, $MSE = 0.12$, $p = .84$, or of reward group, $F(2, 71) = .76$, $MSE = 0.10$, $p = .47$. However, the Item Difficulty × Reward Group interaction was significant, $F(2, 71) = 6.25$, $MSE = 0.77$, $p < .01$, $\eta_p^2 = .15$. Follow-up analyses were conducted to unpack this interaction. A 2 (item difficulty) × 2 (reward group) ANOVA revealed a significant Item Difficulty × Reward group interaction between the high-reward easy group and the high-reward difficult group, $F(1, 47) = 11.69$, $MSE = 1.54$, $p < .001$, $\eta_p^2 = .20$, which indicates that learners selected previously recalled items that were slated with a high reward for study regardless of item difficulty. Next, we compared the high-reward easy group and the constant-reward group. The Item Difficulty × Reward Group interaction approached significance, $F(1, 47) = 3.70$, $MSE = 0.43$, $p = .06$. Finally, the high-reward difficult group was compared with the constant-reward group, and the interaction was not significant, $F(1, 48) = 2.87$, $MSE = 0.36$, $p = .10$.

Self-paced study for unrecalled items. Mean self-paced study time for easy and difficulty items are presented in Table 3. An effect for item difficulty was significant, indicating that participants are spending more time overall studying difficult items, $F(1, 77) = 8.65$, $MSE = 10.47$, $p < .01$, $\eta_p^2 = .10$. The effect for reward group was not significant, $F(2, 77) = 0.41$, $MSE = 11.06$, $p = .66$, η_p^2. Again, the Item Difficulty × Reward Group interaction was significant, $F(2, 77) = 4.05$, $MSE = 42.34$, $p < .05$, $\eta_p^2 = .10$. As in previous analyses, this interaction was explored with several follow-up ANOVAs. We first compared the two high-reward groups. A 2 (item difficulty) × 2 (reward group) ANOVA revealed a significant Item Difficulty × Reward Group interaction, $F(1, 52) = 7.78$, $MSE = 81.23$, $p < .05$, $\eta_p^2 = .10$, which indicated that participants studied high-reward items longer than low reward items. Next, the constant-reward group was compared with the two high-reward groups. The comparison between the constant-reward group and the high-reward difficult group also revealed a significant interaction, $F(1, 49) = 4.73$, $MSE = 39.67$, $p < .05$, $\eta_p^2 = .09$, but the comparison involving the high-reward easy group did not, $F(1, 53) = 0.77$, $MSE = 6.63$, $p = .38$.

Discussion

The goal in Experiment 3 was to ensure that item difficulty would be salient during item selection by explicitly presenting the level of difficulty for each item while participants selected items for restudy. Doing so was expected to highlight variations in item difficulty, so that if participants preferred to use item difficulty to make allocation decisions, information about item difficulty would be readily available. Even under these conditions, however, participants still relied on reward during item selection, as predicted by the ABR model.

Experiment 4

In Experiment 4, our goal was to examine how people may instantiate agendas by exploring factors that are expected to disrupt successful agenda execution. According to the ABR model, agenda execution is a top-down process that requires cognitive resources, so even subtle manipulations that tax these resources may undermine execution. For instance, Thiede and Dunlosky (1999) argued that the format of item presentation for selection could influence the resources learners have available while selecting items for restudy. They compared a simultaneous presentation of items for selection (as used here in Experiments 1–3) with a sequential presentation of items. The latter involved presenting one item at a time and asking participants whether they wanted to select the current item for restudy. As compared with the simultaneous format, the sequential format may require participants to keep extra information active in working memory during item selection. The simultaneous format allows one to externally compare all items, to evaluate which ones meet the criteria of the agenda, so that an appropriate subset can be easily selected for restudy. The sequential format makes comparing items difficult because the current item (along with its reward) would need to be compared with previously presented items that would be available only from memory. Moreover, the simultaneous presentation of both rewards (e.g., 5 versus 1) provides an external reminder of the agenda (e.g., "focus on the unrecalled 5 point items"), which would not be readily available from the sequential array. Thus, the agenda itself may be more externalized under a simultaneous format but would need to be consistently refreshed in working memory under a sequential format. Put differently, the extra demands of the sequential format may diminish cognitive resources in a manner that would disrupt the controlled top-down processing that drives ABR (cf. Barret et al., 2004; Sobel et al., 2007).

Thiede and Dunlosky (1999) demonstrated that the selection format moderated the effect of item difficulty on item selection. When participants were instructed to study the 6 easiest items of a 30-item list, they chose about 6 of the easiest items under a simultaneous format but were more likely to select the majority of the most difficult items under a sequential format. The greater demands of selection under the sequential format presumably caused some participants to forget the initial instructions, and hence their attention was captured by a different aspect of item difficulty when they were selecting items for study. Nevertheless, in the current context, one might intuitively expect that the selection format would not influence item selection, because even under the sequential format, the reward for a pair (5 or 1) was always presented directly above it during selection. In contrast to this intuition, the ABR model predicts that even a salient manipulation, like point reward, will have a diminished influence on selection under the sequential format. To test this prediction, we examined whether the format of items for selection moderated the effects of reward on item selection.

This experiment also allowed us to provide preliminary evidence relevant to evaluating an alternative interpretation of the present effects of reward structure on item selection. According to this alternative, manipulating reward (or test probabilities) directly changes the criterion of mastery for an item. Thus, in the high-reward easy group (in which easy items are slated for a 5-point reward and difficult items are slated for a 1-point reward), the reward would directly influence the criterion of mastery, so that the more difficult items would be treated as already mastered, and hence the easier 5-point items would be selected for study. If so, the discrepancy reduction mechanism or the RPL hypothesis (which includes a criterion for mastery, called the norm of study) could account for the present results. If reward is directly influencing people's criteria of mastery in this manner, then we would not expect the selection format to moderate the effects of reward on item selection because the reward is always presented with an item during selection for study.

Finally, we also had participants perform a reading span (RSPAN) task that taps working-memory capacity (Daneman & Carpenter, 1980; Just & Carpenter, 1992). If ABR is resource demanding, then under the most demanding context in which items are presented sequentially for selection, reward is expected to have an even smaller influence on selection for participants with lower span scores than for those with higher span scores. Although exploratory, if this pattern were obtained, it would provide further evidence that learners must have sufficient resources available to successfully implement ABR.

Method

Participants. Eighty-three participants from Kent State University participated for course credit in Introductory Psychology. Participants were randomly assigned to either the high-reward easy group with a sequential item selection format ($n = 22$), the high-reward difficult group with sequential item selection format ($n = 21$), the high-reward easy group with simultaneous item selection format ($n = 20$), or the high-reward difficult group with a simultaneous item selection format ($n = 20$).

Materials and procedure. The same materials used in Experiments 1 and 2 were used in Experiment 4. The procedure for this experiment was the same as in Experiment 3, with the following exceptions. First, there was no constant-reward group. Second, we ran participants individually, so an RSPAN task could be administered. Participants completed the RSPAN task prior to completing the experimental task. Third, participants did not make EOL judgments in this experiment, and they were not cued to item difficulty because (a) Experiment 3 indicated that doing so did not influence item selection and (b) the main aim in Experiment 4 was to evaluate whether the selection format moderated reward effects. Fourth, and most important, participants selected items for restudy either under a simultaneous format or under a sequential format.

Participants in the simultaneous item selection groups selected items for restudy from a 5×6 array that was identical to the arrays used in Experiments 1 and 2. Participants in sequential item selection groups made their study decisions individually. The cue for each word pair was presented one at a time and the corresponding point value was presented directly above the cue. Participants were prompted to select a pair for restudy by typing a *1*, if they wanted to restudy the pair they were currently viewing, or a *0*, if they did not want to restudy the pair. Self-paced study and final recall were not collected in this experiment because the focal predictions concerned only item selection.

RSPAN task. We used a computer-paced version of the RSPAN task modified from Kane et al. (2004). Participants first were shown either a logical sentence or a nonsensical sentence and then were shown an unrelated letter. Participants read the sentence (e.g., "Wendy studied 1 hr for her math test"), decided whether it made sense, and then studied the letter (e.g., "F"). After the letter was presented, the next sentence–letter dyad appeared onscreen (e.g., "Tom lent Frank money for a ticket to the fish" followed by a "P"). Participants had 4 s to read the sentence and 1 s to study each letter. After the final letter on each trial, a recall cue prompted participants to type the target letters in serial order. The RSPAN task consisted of 15 trials that ranged from three to seven sentence–letter dyads presented in random order. Performance was computed with partial-credit unit scoring in which the mean proportion of correctly recalled letters are aggregated over all trials regardless of set size (for rationale, see Conway et al., 2005).

Results

Preselection recall. A 2 (item difficulty) × 2 (point reward) ANOVA was conducted on preselection recall performance (Table 1); note that selection format was not included as a factor here because this manipulation occurred after preselection recall. Participants recalled more easy items than difficult items, $F(1, 79) = 134.76$, $MSE = 0.02$, $p < .001$, $\eta_p^2 = .63$. The main effect for reward group, $F(3, 79) = 0.10$, $MSE = 0.07$, $p = .96$, and the interaction, $F(3, 79) = 0.65$, $MSE = 0.02$, $p = .59$, were not significant. Preselection recall did not differ as a function of span performance (for details on analyses of span performance, see the section on item selection). Thus, preselection recall conditionalized on span is not presented here but is available from John Dunlosky.

Item selection for unrecalled items. Mean proportion of unrecalled items selected for restudy are presented in Figure 4. A 2 (reward group: high-reward easy vs. high-reward difficult) × 2 (selection format: sequential vs. simultaneous) × 2 (item difficulty: easy vs. difficult) ANOVA revealed no effects for item difficulty, $F(1, 78) = 0.61$, $MSE = 0.09$, $p = .44$, or for reward group, $F(1, 78) = 0.61$, $MSE = 0.03$, $p = .40$. The significant Item Difficulty × Reward Group interaction indicated that participants selected more high-reward items for restudy, $F(1, 78) = 56.78$, $MSE = 4.97$, $p < .001$, $\eta_p^2 = .42$. The effect for selection format approached significance $F(1, 78) = 3.44$, $MSE = 0.16$, $p = .07$. The Item Difficulty × Format interaction was not significant, $F(1, 78) = 0.17$, $MSE = 0.02$, $p = .68$, nor was the Format × Reward Group interaction, $F(1, 78) = 0.02$, $MSE = 0.001$, $p = 90$. However, the three-way interaction was significant, $F(1, 78) = 12.43$, $MSE = 1.09$, $p < .001$, $\eta_p^2 = .14$.

As evident from inspection of Figure 4, participants' preference for selecting high-reward items was weaker when they selected items for restudy under a sequential than simultaneous format. Planned comparisons were consistent with this observation. In particular, a 2 (item reward) × 2 (selection format) ANOVA was first conducted for the high-reward easy groups. The significant interaction, $F(1, 39) = 6.46$, $MSE = 0.42$, $p < .05$, $\eta_p^2 = .14$, indicated that the effect of reward was weaker under the sequential format than under the simultaneous format. The same interaction for the high-reward difficult groups, $F(1, 39) = 6.21$, $MSE = 0.68$, $p < .05$, $\eta_p^2 = .14$, again demonstrated that reward had a smaller influence on item selection under the sequential format.

Relation between span performance and item selection. To allow us to examine the relation between RSPAN performance and item selection, participants were divided into high spans and low spans, from a median split (median = 0.77) on RSPAN scores. Participants with RSPAN scores equal to or above 0.77 were classified as high spans and any participants with scores below 0.77 were classified as low spans. A median split was used in this experiment because our sample was too small to form extreme groups, which has been the preferred method of analyses (e.g., see Conway et al., 2005). Note, however, that a median split would tend to decrease the likelihood of discovering span effects, which would run counter to the predictions from the ABR model.

First, we examined the influence of span on item selection under a simultaneous format. Both high spans and low spans in the high-reward easy group selected a higher proportion of easy items (high span: $M = .71$, $SE = .10$; low span: $M = .83$, $SE = .09$) than difficult items (high spans: $M = .26$, $SE = .07$; low spans: $M = .38$, $SE = .08$), and high spans and low spans in the high-reward difficult group selected a higher proportion of difficult items (high span: $M = .83$, $SE = .06$; low spans: $M = .77$, $SE = .08$) than easy items (high span: $M = .20$, $SE = .07$; low spans: $M = .27$, $SE = .09$). This resulted in a significant Item Difficulty × Reward Group interaction, $F(1, 35) = 52.25$, $MSE = 5.1$, $p < .001$, $\eta_p^2 = .60$. The three-way interaction was not significant, $F(1, 35) = 0.30$, $MSE = 0.03$, $p = .59$. Thus, as expected, under a simultaneous format for item selection, participants preferred to select higher reward items for restudy, regardless of their span scores.

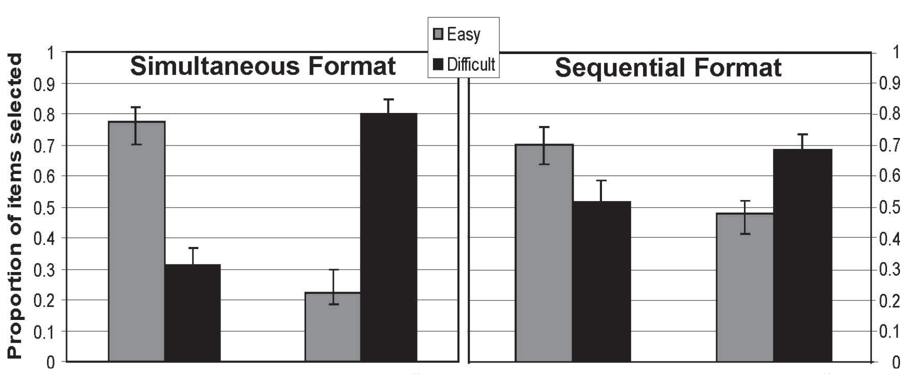


Figure 4. Mean proportion of items selected for restudy (Experiment 4) for items that were not correctly recalled during preselection recall as a function of selection format. Error bars represent standard error of the mean.

Second, and most important, we examined sequential item selection as a function of span performance. Under this more demanding format, the effects of reward were expected to be smaller for low spans than for high spans. As evident from inspection of Figure 5, this prediction was confirmed. The Item Difficulty × Reward Group interaction was significant, $F(1, 39) = 8.48$, $MSE = 0.68$, $p < .01$, $\eta_p^2 = .18$, and the three-way interaction approached significance, $F(1, 39) = 3.10$, $MSE = 0.25$, $p = .09$. Given that we expected this three-way interaction, we conducted a series of follow-ups to explore it further. A 2 (item difficulty) × 2 (reward group) ANOVA was conducted to compare item selection for only high spans. A significant Item Difficulty × Group crossover interaction was obtained, $F(1, 21) = 13.37$, $MSE = 0.94$, $p < .001$, $\eta_p^2 = .39$. The same analysis was conducted for only low spans, and the interaction was not significant, $F(1, 18) = 0.54$, $MSE = 0.05$, $p = .47$.

Item selection for recalled items. No main effects occurred on selection of items that were correctly recalled during preselection recall (Table 2): item difficulty, $F(1, 78) = 0.92$, $MSE = 0.11$, $p = .34$; reward group, $F(1, 78) = 0.01$, $MSE = 0.01$, $p = .91$; and selection format, $F(1, 78) = 0.02$, $MSE = 0.01$, $p = .89$. The Item Difficulty × Group interaction was significant, $F(1, 78) = 25.11$, $MSE = 2.64$, $p < .001$, $\eta_p^2 = .24$, and the three-way interaction was also significant, $F(1, 78) = 11.56$, $MSE = 1.22$, $p < .001$, $\eta_p^2 = .13$: Participants in the simultaneous format groups selected more high-reward items for restudy than did participants in the sequential group. No other interactions were significant. The item selection results for high spans and low spans were consistent with the results described above for the unrecalled items, and the results can be obtained from John Dunlosky.

Discussion

In summary, reward effects on item selection were weaker when participants selected items under a sequential format than when participants selected items under a simultaneous format. Moreover, for lower span participants, reward had no significant influence on their selection when items were presented in the sequential format, whereas high span participants preferred to select items slated for a higher reward, even under the sequential format. If point reward was directly causing participants to change their criterion of mastery for items, selection format should not influence item selection because the reward was always presented with the item. Thus, the present differences (either due to selection format or due to individual differences in span) cannot be explained by assuming that point reward directly influences one's criterion of mastery. By contrast, the ABR model can account for the present effects because in that model, it is assumed that learners must have sufficient resources to maintain and execute their agenda.

General Discussion

The model of ABR claims that learners assess task constraints prior to study and then construct an agenda that aims to efficiently achieve the current task goals within those constraints. In the current experiments, we focused mainly on a single task constraint—the reward structure of the task—and its influence on item selection. In accord with the ABR model, learners were expected to construct an agenda that included their decision criteria for the current task, which would involve prioritizing items for restudy that would potentially maximize reward. Results from Experiments 1–3 were consistent with this prediction: Participants were more likely to select for study items that were more likely to yield higher reward, regardless of item difficulty.

The ABR model assumes that learners strive to allocate study time in an optimal manner (that minimizes study time and maximizes goal achievement), and in Experiments 1 and 2, it was evident that participants' allocation of study led to some expected gains. In particular, as evident from Table 4, gains in performance after self-regulated study were greater for items that participants expected to be more likely to appear on the final test (Experiment 1) and for items that were slated for a higher reward (Experiment 2). Participants obtained labor-and-gain effects because they selected the more highly valued items (Figures 1 and 2) and used more time studying them, too (Table 3). Even though these labor-and-gain effects do indicate that participants' regulation can be effective, we suspect that the methods used in the present experiments limited the potential optimality of participants' self-regulated study. For instance, in Experiments 1 and 2, after a

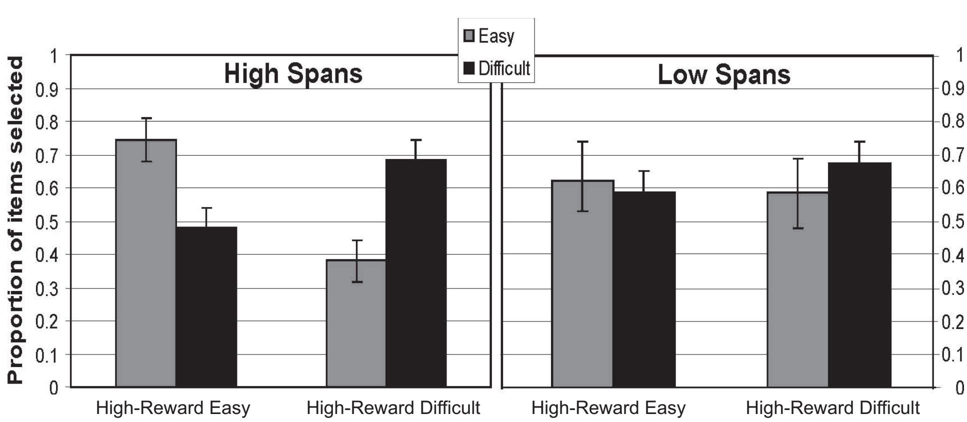


Figure 5. Mean proportion of items selected for restudy under the sequential format (Experiment 4) for items that were not correctly recalled during preselection recall as a function of span score. Error bars represent standard error of the mean.

participant selected an item for restudy, the computer program immediately presented the item for a single restudy trial. Other forms of restudy would be much more potent at enhancing final recall performance. As compared with this single study trial for selected items, participants would likely have done much better if they had multiple opportunities to restudy those items, or even better, if the selected items were retested and restudied across multiple trials (for a review, see Roediger & Karpicke, 2006).

A key point here is that performance gains from self-regulated study will be moderated not only by which items participants select for restudy but also by how the selected items are restudied (for other influential factors, see Son & Sethi, 2006). The ABR model focuses on the former aspect of self-regulation, so the present experiments were designed to evaluate predictions about item selection. Nevertheless, learners may develop agendas about how to reschedule items for restudy, and hence, the ABR model may eventually be extended to account for this aspect of self-regulated learning too. Critical to this endeavor will be a systematic investigation of how learners decide to schedule restudy for selected items. So, if learners are given the opportunity both to select items for restudy and to decide how to schedule restudy for the selected items, will they select optimal practice schedules and strategies? Some investigators are just now beginning to offer answers to this question. Benjamin and Bird (2005) and Son (2004) reported that college students chose spaced practice more often than massed practice to restudy items, and doing so boosted their level of final test performance. Kornell and Bjork (2007) found that as college students' memory for a to-be-learned list increased, they began testing (versus restudying) their memory more often, which may be an effective strategy for both monitoring and learning lists of words. This evidence suggests that learners may make relatively good decisions about how to schedule practice trials, although other evidence indicates that learners will not make optimal decisions under some circumstances (e.g., Benjamin, 2007; Kornell & Bjork, 2007). Perhaps our most important point for now, however, is that regardless of whether learners achieve optimality as they regulate study, their construction and use of agendas has a major—and underinvestigated—influence on study-time allocation.

Current State of Theories of Study-Time Allocation

In most theories of self-regulated study, it is assumed that item difficulty plays a key role in influencing learners' decisions to allocate study. The results of the present experiments demonstrate that this influence is not universal. Though item difficulty may influence decisions about how to allocate study under many conditions (for a review, see Son & Metcalfe, 2000), other task constraints, such as potential reward for learning an item, can also affect their decisions. Reward for learning may have a particularly important impact on item selection in naturalistic settings. Learners outside of the laboratory are often faced with decisions about what information is important for them to commit to memory. During self-regulated study, this decision is often made to achieve some goal, such as earning a specific grade on an examination. Attaining such goals requires learners to accurately identify what material they will need to remember and to direct cognitive resources toward learning just this material. This behavior reflects the fact that learners are motivated to prioritize information for study that will maximize test performance (Castel, 2007).

The outcomes from the present experiments provide important constraints for theories of self-regulated study in general and uncover potential limitations of two of the most extensively investigated theories in the field. In particular, the discrepancy reduction theory predicts that learners will always select the most difficult items for study (Dunlosky & Hertzog, 1998), and the RPL theory predicts that learners will select the easiest unlearned items for study (Metcalfe, 2002). With these theories, it is difficult to explain the crossover interactions presented in Figures 1 through 4. One way they could account for these effects is by assuming that reward structure directly influences the criteria of mastery for items. Even with this assumption, however, the theories will have difficulties accounting for the diminished effects of reward on item selection as a function of selection format (Experiment 4). By contrast, the ABR model provides an intuitively plausible and parsimonious explanation for the crossover interactions in Experiments 1 through 3 and the diminished reward effects in Experiment 4. Note, however, that even though we competitively evaluated predictions from the ABR model and the item-difficulty based theories of self-regulated study, we acknowledge that the latter theories were not developed to account for the effects of reward structure. Thus, our main conclusion is not that discrepancy reduction or one's RPL do not play a role in self-regulated study but that a general model will need to go beyond these theories to account for all facets of learners' self-regulation.

The ABR model differs from other theories of self-regulated study because it proposes that agendas can dominate regulation, whereas the aforementioned theories focus on how monitoring item difficulty drives regulation. We are not arguing that learners never monitor item difficulty (or their on-going learning) in the service of regulating study, but for the ABR model, monitoring item difficulty is used in the service of fulfilling one's agenda. For instance, when students have very little time to study, they may develop an agenda that involves identifying the easiest material to study first; and in this case, executing that agenda would involve monitoring item difficulty (Dunlosky & Thiede, 2004). In many cases, however, the agenda will not require monitoring item difficulty (such as in the present experiments) and instead will be selected on other task relevant factors, such as item reward or the likelihood of an item appearing on a test.

Note, also, that we designed these experiments to evaluate whether reward structure could dominate participants' item selection: The reward structures we chose were meant to potentially produce large effects, and reward structure covaried with item difficulty within each group. Given the latter, we could not evaluate the degree to which learners' item selection was jointly influenced by reward structure and item difficulty. Important trade-offs will likely occur; for instance, as the difference in point reward for items becomes smaller (e.g., 1 versus 2 instead of 1 versus 5), participants may shift toward using item difficulty to allocate study time. Likewise, as point reward becomes even more disparate (e.g., 1 point versus 100 points), participants may never select lower valued items for restudy (whereas they did in the present studies; see, e.g., Figure 1). Exploring how various sources of information mutually drive allocation of study time is an important avenue for future research. To further guide such research, we compare the ABR model with two theories that have received attention in the literature: the hierarchical model (Thiede & Dunlosky, 1999) and the RPL theory (Metcalfe, 2002).

Hierarchical model of self-regulated study. As discussed in the introduction, the hierarchical model (Thiede & Dunlosky, 1999) is similar to the ABR model because in both theories, it is assumed that learners use agendas when regulating study. The hierarchical model consists of a superordinate level, in which learners construct a plan that indicates which items the learner needs to study, and a subordinate level in which study is controlled by the discrepancy-reduction mechanism. According to this model, in situations in which learners do not use a plan, learners' item selection is driven solely by discrepancy reduction (Dunlosky & Thiede, 2004).

In the hierarchical model, study is not dominated by planning because self-paced study times of selected items are expected to be driven solely by discrepancy reduction. In the current experiments, even for items that were selected for study, self-paced study time was influenced by the reward structure of the task (Table 3), which disconfirms this core prediction from the hierarchical model. Moreover, evidence from previous experiments suggests that discrepancy reduction may have a limited role in regulating study time (Metcalfe, 2002; Metcalfe & Kornell, 2005). For these reasons, the hierarchical model—in which planning and discrepancy reduction are inextricably linked—has limited explanatory power and no longer seems a viable model as it was originally proposed. By contrast, discrepancy reduction does not drive ABR. Instead, the agenda controls regulation and learners' use of monitoring to make strategic decisions about how to allocate study. As noted earlier, even under ABR, item difficulty can influence item selection, such as when a learner's agenda is constructed in an attempt to learn material strategically, using item difficulty.

More generally, as compared with the hierarchical model, the ABR model emphasizes the critical role of agendas in driving regulation, but it does not indicate that monitoring item difficulty must be used in the same fashion—that is, focusing on the most difficult items—across all learner goals and task constraints. Item difficulty may also be used in other ways to guide self-paced study time: Learners may spend more time studying items within their RPL (Son & Metcalfe, 2000) or may monitor ongoing progress while studying (Dunlosky & Thiede, 1998; Metcalfe & Kornell, 2005). Any of these mechanisms are viable additions to ABR. Thus, investigating the interplay between ABR and nonagenda based mechanisms that are driven by item difficulty is an important avenue for future research.

RPL theory. The RPL theory has received much empirical support in the literature on self-regulated study (for a review, see Son & Kornell, 2008). Consider evidence from an exemplary experiment that is consistent with predictions from the RPL theory. Metcalfe (2002, Experiment 1) had students study English–Spanish vocabulary. Some pairs were objectively easy to learn (e.g., *fantastic–fantastico*), some pairs were intermediate (e.g., *husband–marido*), and some were difficult to learn (e.g., *closet–guardaropa*). During a given trial, the stimulus from one pair from each difficulty level was presented simultaneously on a computer screen, and to study an item, participants clicked on a question mark that appeared below a stimulus (and then the Spanish translation would appear). The order of presenting items on the screen increased in difficulty from left to right, and participants were instructed that the item on the far left was the easiest to learn, the item in the middle had an intermediate difficulty, and the item on the far right was the most difficult to learn. It is important to note that on any given trial, the pairs could be studied for a total of 5 s, 15 s, or 1 min. According to RPL theory, with very little time (5 s), participants' RPL would more likely include the easy-to-intermediate items because the difficult ones would be too difficult to learn in that time frame. As the available study time increased, however, the RPL would include more of the difficult items. Thus, as times increased, students should shift from focusing on mainly the easy-to-intermediate items to studying the more difficult ones. This prediction was fully supported by the data (in this experiment and four others), which provided impressive support for the RPL theory.

A question arises as to the nature of the RPL. In the present case (Metcalfe, 2002), one answer is that the behavioral marker for RPL regulation is indicative of ABR. Learners in Metcalfe's (2002) experiments may realize that the limited amount of time available to study all the items would result in little benefit if difficult items were selected for study. That is, on any given trial, a participant would alter his or her agenda on how to maximize performance given the current task constraints (i.e., the amount of study time available). If so, at least in some situations, agenda construction and execution would underlie the adoption of an RPL scheme for regulation. The idea is that depending on the task constraints, learners may adopt an agenda that involves selecting first for restudy items that are included in the RPL. In such cases, RPL would represent a specific instantiation of ABR. However, under other task constraints (e.g., the present experiments), learners' agenda will involve selecting items on the basis of non-RPL criteria. In this manner, regulation of item selection is highly flexible and can be adjusted in an attempt to meet varying goals and task constraints. An alternative possibility, however, is that although some regulation is agenda driven, in other circumstances, allocation of time to a RPL is driven by another mechanism. That is, RPL regulation may not be an instantiation of ABR, but instead may be a prepotent response to regulating study time under some task conditions.

Conclusions

Previous research in which self-regulated study was examined has not yet explored many factors of the learning context that may influence study decisions and instead has largely emphasized the core role of item difficulty in influencing item selection. This focus on the effects of item difficulty stems from the view that learners' regulation of study reflects monitoring and control relations that are sensitive to processing item difficulty and ongoing learning (Nelson & Leonesio, 1988). It is not our goal to dispute the claim that monitoring can drive study behavior, because it does (Metcalfe & Finn, 2008). Rather, our current goal is to emphasize that learners can use monitoring and control processes in a highly flexible manner. Learners can use both processes to carry out an agenda that is aimed at meeting their goals in an efficient manner. And as demonstrated in the current experiments, monitoring item difficulty may be overridden as a basis for regulation when agendas are constructed to meet opposing task goals.

References

Barret, L. F., Tugade, M. M., & Engle, R. W. (2004). Individual differences in working memory capacity and dual-process theories of the mind. *Psychological Bulletin, 130,* 553–573.

Benjamin, A. S. (2007). 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. (2005). Metacognitive control of the spacing of study repetitions. *Journal of Memory and Language, 55,* 126–137.

Carroll, M., Nelson, T. O., & Kirwan, A. (1997). Tradeoff of semantic relatedness and degree of overlearning: Differential effects on metamemory and on long-term retention. *Acta Psychologica, 95,* 239–253.

Carver, C. S., & Scheier, M. F. (2000). On the structure of behavioral self-regulation. In M. Boekaerts, P. R. Pintrich, & M. Zeidner (Eds.), *Handbook of self-regulation* (pp. 41–84). New York: Academic Press.

Castel, A. D. (2005). Memory for grocery prices in younger and older adults: The role of schematic support. *Psychology and Aging, 20,* 718–721.

Castel, A. D. (2007). The adaptive and strategic use of memory by older adults: Evaluative processing and value-directed remembering. In A. S. Benjamin & B. H. Ross (Eds.), *The psychology of learning and motivation: Skill and strategy in memory use.* (Vol. 48, pp. 225–270). London: Academic Press.

Castel, A. D., Benjamin, A. S., Craik, F. I. M., & Watkins, M. J. (2002). The effects of aging on selectivity and control in short-term recall. *Memory & Cognition, 30,* 1078–1085.

Castel, A. D., Farb, N. A. S., & Craik, F. I. M. (2007). Memory for general and specific value information in younger and older adults: Measuring the limits of strategic control. *Memory & Cognition, 35,* 689–700.

Connor, L. T., Dunlosky, J., & Hertzog, C. (1997). Age-related differences in absolute but not relative metamemory accuracy. *Psychology and Aging, 12,* 50–71.

Conway, A. R. A., Kane, M. J., Bunting, M. F., Hambrick, D. Z., Wilhelm, O., & Engle, R. W. (2005). Working memory span tasks: A methodological review and user's guide. *Psychonomic Bulletin & Review, 12,* 769–786.

Daneman, M., & Carpenter, P. A. (1980). Individual differences in working memory and reading. *Journal of Verbal Learning and Verbal Behavior, 19,* 450–466.

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

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.

Dunlosky, J., Hertzog, C., Kennedy, M. R. F., & Thiede, K. W. (2005). The self-monitoring approach for effective learning. *International Journal of Cognitive Technology, 10,* 4–11.

Dunlosky, J., & Matvey, G. (2001). Empirical analysis of the intrinsic-extrinsic distinction of judgments of learning (JOLs): Effects of relatedness and serial position on JOLs. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 27,* 1180–1191.

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

Dunlosky, J., & Thiede, K. W. (2004). Causes and constraints of the shift-to-easier-materials effect in control of study. *Memory & Cognition, 32,* 779–788.

Eysenck, M. W., & Eysenck, M. C. (1982). Effects of incentive on cued recall. *The Quarterly Journal of Experimental Psychology: Human Experimental Psychology, 34,* 489–498.

Greene, J. A., & Azevedo, R. (2007). A theoretical review of Winne and Hadwin's model of self-regulated learning: New perspectives and directions. *Review of Educational Research, 77,* 334–372.

Hertzog, C., Kidder, D. P., Powell-Moman, A., & Dunlosky, J. (2002). Aging and monitoring associative learning: Is monitoring accuracy spared or impaired? *Psychology and Aging, 17,* 209–225.

Johnson, M. K., Hashtroudi, S., & Lindsay, D. S. (1993). Source monitoring. *Psychological Bulletin, 114,* 3–28.

Just, M. A., & Carpenter, P. A. (1992). A capacity theory of comprehension: Individual differences in working memory. *Psychological Review, 99,* 122–149.

Kane, M. J., Hambrick, D. Z., Tuholski, S. W., Wilhelm, O., Payne, T. W., & Engle, R. W. (2004). The generality of working memory capacity: A latent-variable approach to verbal and visuospatial memory span and reasoning. *Journal of Experimental Psychology: General, 133,* 189–217.

Koriat, A. (1997). Monitoring one's own knowledge during study: A cue-utilization approach to judgments of learning. *Journal of Experimental Psychology: General, 126,* 349–370.

Koriat, A., Ma'ayan, H., & Nussinson, R. (2006). The intricate relationship between monitoring and control in metacognition: Lessons for the cause-and-effect relation between subjective experience and behavior. *Journal of Experimental Psychology: General, 135,* 36–69.

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

Kornell, N., & Bjork, R. A. (2008). Optimizing self-regulated study: The benefits—and costs—of dropping flashcards. *Memory, 16,* 125–136.

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

Loftus, G. R., & Wickens, T. D. (1970). The effects of incentive on storage and retrieval processes. *Journal of Experimental Psychology, 85,* 141–147.

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

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 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.

Mitchell, K. J., Raye, C. L., McGuire, J. T., Frankel, H., Greene, E. J., & Johnson, M. K. (2008). Neuroimaging evidence for agenda-dependent monitoring of different features during short-term source memory tests. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 34,* 780–790.

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., & Leonesio, 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.

Pintrich, P. R. (2000). The role of goal orientation in self-regulated learning. In M. Boekaerts, P. R. Pintrich, & M. Zeidner (Eds.), *Handbook of self-regulation.* (pp. 451–502). New York: Academic Press.

Rabinowitz, J. C., Ackerman, B. P., Craik, F. I. M., & Hinchley, J. L. (1982). Aging and metamemory: The roles of relatedness and imagery. *Journal of Gerontology, 37,* 688–695.

Reder, L. M. (1988). Strategic control of retrieval strategies. In G. Bower (Ed.), *The psychology of learning and motivation* (Vol. 22, pp. 227–259). New York: Academic Press.

Roediger, H. L., III, & Karpicke, J. D. (2006). The power of testing memory: Basic research and implications for educational practice. *Perspectives on Psychological Science, 1,* 181–210.

Schneider, W., Eschman, A., & Zuccolotto, A. (2001). *E-prime user's guide.* Pittsburgh, PA: Psychology Software Tools.

Schunn, C. D., & Reder, L. M. (1998). Strategy adaptivity and individual differences. In D. L. Medin (Ed.), *The psychology of learning and motivation* (pp. 115–154). York, PA: Academic Press.

Sobel, K. V., Gerrie, M. P., Poole, B. J., & Kane, M. J. (2007). Individual differences in working memory capacity and visual search: The roles of top-down and bottom-up processing. *Psychonomic Bulletin & Review, 14,* 840–846.

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

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

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.

Son, L. K., & Sethi, R. (2006). Metacognitive control and optimal learning. *Cognitive Science, 30,* 759–774.

Souchay, C., & Isingrini, M. (2004). Age-related differences in the relation between monitoring and control of learning. *Experimental Aging Research, 30,* 179–193.

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

Thiede, K. W., & Dunlosky, J. (1999). Toward a general model of self-paced 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.

Weiner, B. (1966). Effects of motivation on availability and retrieval of memory traces. *Psychological Bulletin, 65,* 24–37.

Winne, P. H., & Hadwin, A. F. (1998). Studying as self-regulated learning. In D. J. Hacker, J. Dunlosky, & A. C. Graesser, (Eds), *Metacognition in educational theory and practice.* (pp. 277–304). Hillsdale, NJ: LEA.

Received August 26, 2008
Revision received February 20, 2009
Accepted February 23, 2009 ■