Available online at www.sciencedirect.com

SCIENCE DIRECT®

Journal of Memory and Language 52 (2005) 463–477

Journal of Memory and Language

www.elsevier.com/locate/jml

A Region of Proximal Learning model of study time allocation ☆

Janet Metcalfe*, Nate Kornell

Department of Psychology, Columbia University in the City of New York, New York, NY 10027, USA

Received 1 September 2004; revision received 6 December 2004
Available online 2 February 2005

Abstract

A Region of Proximal Learning model is proposed emphasizing two components to people's study time allocation, controlled by different metacognitive indices. The first component is *choice*, which is further segmented into two stages: (1) a decision of whether to study or not and (2) the order of priority of items chosen. If the people's Judgments of Learning (JOLs) are sufficiently high that they believe they know the items already, they will choose to not study. If they do choose to study, the order is from that which they believe is almost known to that which they believe is more difficult (high JOL to low JOL). The second component is *perseverance*, with emphasis on the rule for stopping studying once study has begun on an item. We propose that people use a previously unexplored process-oriented metacognitive marker: their judgments of the *rate* of learning (jROLs), to decide when to stop. When learning is proceeding quickly and the jROL value is high they continue studying. When the jROL approaches zero, and their subjective assessment indicates learning is at a standstill they stop. The extant literature bearing on this model is reviewed, and eight new experiments, all of which support the model, are presented.

Keywords: Metacognition; Study-time allocation; JOL; jROL; Model

The manner in which our metacognitions enable us to shape and modulate our own knowledge acquisition and impact upon the dynamics of what and how we learn is central to human cognition. At the same time as progress continues in our understanding the conditions, constraints, and mechanisms underlying people's metacognitions (Dunlosky & Nelson, 1992; Hertzog, Dunlosky, Robinson, & Kidder, 2003; Kimball & Metcalfe, 2003; Koriat, 1993, 1994; Koriat & Goldsmith, 1996; Koriat & Levy-Sadot, 2001; Koriat, Sheffer, & Ma'ayan, 2002; Maki & Berry, 1984; Metcalfe, 1993a, 1993b, 1998; Metcalfe, Schwartz, & Joaquim, 1993; Metcalfe & Shimamura, 1994; Miner & Reder, 1994; Nelson & Narens, 1990, 1994; Reder, 1988; Schwartz & Metcalfe, 1992, 1994; Schwartz & Smith, 1997; Spellman & Bjork, 1992; Thiede, 1999), we are witnessing a shift in interest towards the consequences of these metacognitions (Mazzoni & Cornoldi, 1993; Metcalfe, 2002; Metcalfe & Kornell, 2003; Reder & Ritter, 1992; Thiede & Dunlosky, 1999). But, although there has been an increasing focus on how people use metacognitions to alter their study behavior and exert control over their

☆ This research was supported by National Institute of Mental Health Grant MH60637. We thank Lisa Son and Bridgid Finn for their help and comments.
* Corresponding author.
E-mail address: Metcalfe@psych.columbia.edu (J. Metcalfe).

0749-596X/$ - see front matter
doi:10.1016/j.jml.2004.12.001

own future learning and memory, no coherent answer concerning how metacognitions guide our study behavior has yet emerged. Indeed, the dominant answer, the Discrepancy Reduction model (Dunlosky & Hertzog, 1998; and see Dunlosky & Thiede, 1998; Hyland, 1988; Koriat & Goldsmith, 1996; Le Ny, Denhiere, & Le Taillanter, 1972; Lord & Hanges, 1987; Nelson & Narens, 1990; Powers, 1973; Thiede, Anderson, & Therriault, 2003; Thiede & Dunlosky, 1999) appears to be unable account for recent data. We will propose an alternative—a Region of Proximal Learning model—and show how it can both account for the data that gave rise to the Discrepancy Reduction model in the first place, and also predict the new data. In addition, the Region of Proximal Learning model is not trapped by the paradox apparent in the Discrepancy Reduction model: if people study appropriately, as given by that model, they will spend an inordinate, perhaps an unlimited, amount of time trying to learn items that may be unlearnable. They are doomed to 'labor in vain.'

The notion that people have a region in which learning can be optimized, and that they choose to study within it, is intuitive. It also squares with theories of human learning proposed by such distinguished researchers as Atkinson (1972), Berlyne (1978), Piaget (1952), Vygotsky (1987), and Hebb (1949). However, the heuristics that people use to hone in on this region have not been delineated. We outline, here, the underpinnings of a metacognitively guided Region of Proximal Learning model, outlining some of these most important heuristics. First, though, we briefly summarize the alternative Discrepancy Reduction model.

In the Discrepancy Reduction model, and related theoretical positions, such as the 'monitoring-affects-control' proposal of Nelson and Leonesio (1988), people are thought to focus their study on those materials that either are, or are thought to be, most difficult. It is notable that this idea is similar to that used in the popular parallel distributed processing models (Rumelhart & McClelland, 1986) which attack the largest errors first. Negative correlations between choice or study time and judged difficulty are taken to be data favoring the Discrepancy Reduction model. According to the model, people are thought to continue to study until they reach an internal criterion of learning. As Thiede and Dunlosky (1999) put it: "An item will continue to be studied (either through selection or through continued allocation of study time) until the person's perceived degree of learning meets or exceeds the norm of study" (p. 1024). Reaching this criterion of learning takes longer the more difficult are the to-be-learned items.

There are four implications of this model: first, it predicts a negative correlation between study time allocation and judgments of learning: low judgments of learning—indicating that the individual does not know the item—should be associated with both priority of choice and longer study. Second, such a negative correlation indicates appropriate metacognitively guided study time allocation, and the larger the negative correlation, according to this model, the better the metacognitively guided control. Third, like the Region of Proximal Learning model, that will be detailed shortly, and perhaps other models as well, given that study-time allocation is assumed to be appropriately controlled by judgments of learning, it is essential that people be able to make accurate JOLs. If their judgments are faulty, then they should study inappropriately, resulting in poor learning. Fourth, insofar as the rule used to stop studying is attainment of an internal criterion of learning, if people held to this model, and an item were unlearnable, they could, in principle, study for an unlimited amount of time (the labor in vain paradox). We next present the proposed alternative.

The Region of Proximal Learning model

In the Region of Proximal Learning model, there are two separable components to human study time allocation: choice and perseverance. It is important to distinguish between them insofar as they are guided by different metacognitive markers. In the choice stage people decide which items they will study, and in which order. In the perseverance stage the question is: how long, once they have started studying a particular item should they continue before switching to another item? Notice that the perseverance question, as we frame it, is not whether or not the goal has been attained, but rather whether study, at the present time, is having sufficiently beneficial results. This change in perspective will allow the model to cut through the labor-in-vain paradox.

Choice. The choice stage can be divided into two substages. The first is to determine which items are candidates for study, or equivalently, which items people need not bother studying. It is proposed that if people assess that they already know an item they will choose not to study it. The second substage involves order determination of those items that will be studied. When determination of order is encouraged or allowed by the experimental situation, the model proposes that people will opt to study the easiest as yet unlearned items first (items that are 'easy pickings'), turning to more difficult items only later. Of course, if all of the items must be learned and time is unlimited, people may choose all of the items, but we posit that they will still tend to prioritize from easy to difficult if they can.

This two-part choice heuristic is not inconsistent with negative correlations between JOLs and choice order. However, if a negative correlation is observed, it occurs entirely because of the first stage—because people decline to study those very high JOLs that they believe they know already. The second stage should produce a positive, not a negative, correlation. The magnitude of the negative correlation between JOL and study choice does not indicate goodness of metacognitive control in this model, as it does in the Discrepancy Reduction model. Indeed, depending on the proportion of items already known, the correlation between JOL and *appropriate* study choice might be negative, zero, or positive, in this model. This is not to say that one cannot assess whether people are exerting good metacognitive control: this correlation should behave in a principled way. It should be negative when many items are known (or thought to be known), just because people should decline to study the already known items, and this fact, in itself, will produce a negative correlation. However, when few or no items are known to the participant the correlation should be positive—reflecting the fact that choice is proposed to go in the order easiest to hardest, among those items that are not already learned.

Perseverance. We propose that people base the decision to stop on the perceived *rate* at which learning is proceeding. When they perceive that they are learning at a rapid rate, they continue. When they perceive that they are no longer taking in information—that learning is going nowhere—then they stop studying a given item and switch to another. We call this newly posited metacognition a *judgment of the Rate of Learning* (jROL), and emphasize that it refers to the monitoring of an active process—the speed of information intake—rather than to a static state of knowledge. The jROL is the derivative of an online JOL function over time. People continue studying when they have a high jROL (they feel themselves 'on a ROL'); they stop when their jROLs approach zero, or some low criterion value.

Although the jROL hypothesis may sound odd at first, and although there are no data (other than those we will present shortly) that bear on it, in other areas of study—which share an analogous problem of determining when an animal should stop doing what they are doing and turn to something else—similar ideas have emerged. For example, in foraging models, the stop rule is to discontinue feeding in a particular location once the rate of food intake at that patch declines (Stephens & Krebs, 1986). That the perceived *rate* of learning controls whether people persist in or defer from further studying is consistent with the idea that people act as information foragers (see, Pirolli & Card, 1999; Weber, Shafir, & Blais, 2004). Similarly, the idea of diminishing returns is common in economic models, and is a reason for switching (or selling). Interestingly, Dunlosky and Thiede (1998), in a paper interpreted as providing basic support for the discrepancy reduction view, and the stop rule therein (but which also underlined some limitations of that view) mentioned, in the penultimate paragraph, an alternative possibility in which "the stopping rule may be a function of perceived rate of learning" (p. 54). This alternative is just like the jROL idea, proposed here. They further noted that this possibility "provides a plausible alternative to the kinds of discrepancy-reduction model that dominate theory of metacognitive control" (p. 55). Finally, Carver and Scheier (1990) also point to the *rate* of discrepancy reduction as a major determinant of people's affect.

jROLs may approach zero (or some other, participant defined, and/or situationally modifiable stopping criterion) for different reasons: (1) once an item is learned; insofar as no further learning is possible, the jROL goes automatically to zero. This is an obvious case. (2) When as much learning as can occur, immediately, has already occurred, even though a subsequent test may show that the item is not yet really learned, the jROL approaches zero. This case is interesting. If learning at time t has reached an asymptote such that further immediate study is having no effect then the model suggests that further immediate practice (i.e., *massed* practice) is not of value and the person should and will stop studying. However, this does not necessarily mean that the person should not return to the item at some later time, when, under new conditions, the jROLs may be non-zero. Spaced rather than massed practice is indicated for such items. (3) A final case of close to zero jROLs comes with very difficult materials on which the person is making no headway. The model, through the jROL stopping rule, provides a limit to people's 'labor in vain' for such items.

The JOL that should be observable when the person stops studying in (1) should be high—probably 100%. The JOLs upon stopping, in (2) might also be high, but spuriously so—these may be cases where an immediate JOL would be high but a delayed JOL on the same item would be lower. However, JOLs might also be low in this case, if the learning process is, for any reason, blocked at the current moment. The JOLs for (3) will be low, often very low. Thus it is not the absolute JOL that determines stopping, but rather the jROL, which need not, itself, be correlated with the JOL.

Although the jROL is the derivative of the JOL function over time, and could be computed from repeated sampling (and comparison) of JOLs, the phenomenology of the jROL relates more closely to feelings of interest and engagement (when the jROLs are high) or to boredom or stagnation (when the jROLs are low) than to any experience of performing a computation. This concept captures the frustration that people feel when they seem to be getting nowhere, and their discouragement with further study under these conditions. It also relates to the studies of attention to degrees of novelty that Berlyne (1978) described in his early investigations of curiosity and interest. He found that the visual patterns that provoked people's interest and engaged their attention (as measured by duration of gaze and by pupillary dilation) were those that were more challenging than those they could easily understand—not perfectly symmetric, but they were not extremely bizarre, to the point of being incomprehensible. We would argue that the challenging patterns allowed a rapid rate of learning, while the bizarre patterns did not allow learning and, therefore, did not engage people's attention. They shifted their gaze away from these items quickly.

Factors such as the reward structures of the task, the amount of time available, the difficulty of the materials, and the self-expectations and drives of the individual alter the parameter value of the stopping jROL. This value will be higher when people have little time than when time is unlimited, closer to zero (lower) on tasks that people care about and which they feel are central for their self concept, than on tasks they think are irrelevant and unimportant; higher when the alternative materials they could turn to rather than persisting on the present item are many and relatively easy, than when they are few and difficult. Regardless of the fact that one can manipulate the stopping jROL value, we argue that low jROLs are inherently aversive, and that the strong feeling of learning that is the phenomenological accompaniment of a high jROL is, itself, pleasurable and motivating for humans.

Data concerning choice

Do people choose to decline study for the items that they know they know?

A number of studies, several going back three decades (Atkinson, 1972; Masur, McIntyre, & Flavell, 1973), affirm people's capability to decline study of already-known items. We have conducted several demonstration experiments focusing on this topic that will be presented below. Masur et al. (1973) found convincing support for a choice strategy in which the already-known items are spurned. In their study with children of various ages, even very young children showed a tendency to choose the items they did not know for study and to eliminate those they did, and this characteristic became quite marked in college students. Grade 1 children were less likely to make this choice than were older children, but that seemed to result from the difficulty they experienced in knowing what they did not know. Thus, they sometimes chose, incorrectly, to not study items that they did not yet know—a tendency that disappeared by Grade 3.

In a second study bearing on this issue, Atkinson (1972) told participants that "their trial-to-trial selection of items should be done with the aim of mastering the total set of vocabulary items" and "that it was best to test and study on words they did not know rather than on ones already mastered" (Atkinson, 1972, p. 125). Although he did not report data in terms of whether or not the items chosen in this so-called 'self-selection' condition had already been learned or not, he did compare the results of the self-selection condition to those in which a computer algorithm selected for people the items for study based on their performance history, which allowed determination of whether items were or were not learned. Performance in these two groups was very similar, at every stage of learning, with one group showing slightly higher performance on one trial and the other group reversing this on the next trial.[1] These results suggest that people were able to conform to the instructions to eliminate from further study those items that were already known. Similarly, Cull and Zechmeister (1994) explicitly told participants to select items for study if and only if they were unlearned. The items they chose to study, given these instructions at each choice, were the lower JOLs.

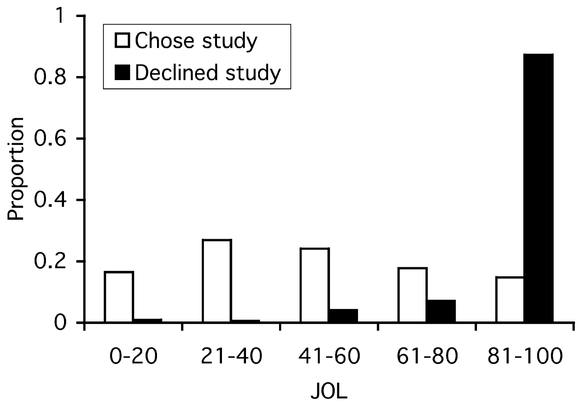

Fig. 1. The distributions of JOLs for items on which people chose to study and for items on which people declined further study. From Son (2004).

In an experiment that was investigating the relation of people's JOLs to their preferences of whether to mass or space practice, Son (2004) included an option of allowing people to decline further study entirely. Fig. 1 shows the distribution of people's JOLs conditional upon their having chosen to study (i.e., collapsed over the massed or spaced choices), or conditional upon their having chosen to decline study. The data were sorted into the JOL intervals based on the raw JOL ratings given in the original data. (We thank Son for providing

[1] The data from these two groups, in Atkinson's (1972) experiment, were quite different from a group given a random assignment of items (including study of already known items) which resulted in much poorer performance. They also differed from a group given only those items that were close to being learned, in which the items that were already learned as well as those that were very far from being learned, as determined by the computer algorithm, were eliminated. This latter condition showed superior performance.

her data set for this analysis.) As can be seen, when people declined further study, their JOLs were nearly always very high. In contrast, JOLs were fairly evenly distributed when people chose to study.

In summary, then, although the literature on the first stage of the choice process proposed by the Region of Proximal Learning model is clear, it is not abundant. To further elaborate on people's choices to study or not study, as a function of whether or not they know the items, we conducted several experiments.

Experiment 1

In the first experiment, with 24 Columbia University students, we used Nelson and Narens' (1980) general-information-question pool, pruned to eliminate those questions that were out of date. A typical question was "Who was the first Prime Minister of Canada?" The computer presented a random selection of such questions, one at a time, without the answers, to participants. They were asked to choose exactly half of the questions for further study of both the question and the answer. They were told that "your goal is to do as well as you can on the final test." On each of the 6 trials, each consisting of a total of 32 questions, the computer displayed a counter, showing dynamically how many of the allowable choices had accumulated in the 'study' and 'don't study' categories. The participant decided whether or not to study each item, one at a time, by clicking on either 'study' or 'don't study' as each item was presented. When the participant had exhausted one choice option, the other option was the only possible response. These trials, where the choices were constrained rather than free, were not included in any of the analyses that follow, either for this experiment, or for any of the subsequent experiments that used a similar design. (One consequence of this is that if participants select 'study' more often than 'don't study' until they have exhausted that choice option, the proportion of 'study' trials will be greater than .5.)

In a survey at the end of the experiment, 20 of 24 participants reported selecting for study the items to which they did not know the answers. None reported selecting known items. The items that were selected for study were more difficult (180.90) than were those not selected (138.64), based on Nelson and Narens' (1980) norms, $t(23) = 7.88$, $p < .0001$, Effect Size = .73. This pattern held for 22 of the 24 participants. People, apparently, had selected the items they did not know, both according to self report, and to the normative data.

It appears, then, with semantic memory general information questions on which it is known, from the considerable past research on this particular item pool, that people have very good metacognitive accuracy, people choose to study those items that they know they do not know, and will decline to study those that they do. Items in semantic memory, such as the general information questions in this first experiment, might be expected to be easy to monitor. In the experiments we report below, we sought to determine whether such a clear pattern also emerged with new learning. We conducted three new-learning experiments in which in the first phase participants were allowed to study only half of the to-be-learned materials (so that the other half was necessarily unlearned). Following study, participants were given the cues for all of the pairs and asked to select half of the items for (re)study. They were then allowed to (re)study, and were tested.

Experiment 2

In this experiment, participants were 24 Columbia University students who participated for either course credit or for pay. They studied highly associated pairs of words with cue-to-target associability scores within the .050–.054 range, based on norms published by Nelson, McEvoy, and Schreiber (1998). These were pairs such as "well-done," which were easy to learn quickly, but which had only a low probability of being correctly produced without any exposure, that is, by guessing. We used these pairs because we thought that a single relatively brief observation of the target would be sufficient to allow full learning—making it highly likely that people would know that they knew these items.

Each participant was tested on four lists each of which consisted of 24 item-pairs. Participants were only presented with half of the pairs in each list during the study phase, however. On two of the lists, the pairs that were presented were shown six times, each for 2.5 s in a spaced manner. On the other two lists the presented pairs were shown only once for 2.5 s. Order of the repetition condition (6 or 1) was counterbalanced across participants. The item-pairs were presented for study, with only one pair showing at a time, on Macintosh computers. Participants were informed that although they only would be shown half of the to-be-remembered items in the initial study phase, they would be shown all of the test cues, in the choice phase, and that they would be tested on the entire set of 24.

The participants were then presented with each of the 24 cues (12 of which were new), one at a time, in a random order, and asked to choose whether they did or did not want to study that cue-target pair for an upcoming test in which they would be asked the responses to all cues. They were told: "You will only be allowed to select half of the pairs for additional study. This makes it very important that you make your selections carefully—keep in mind that you will be tested on all of the pairs, not just the ones you select, and your goal is to do as well as you can on the final test." As in Experiment 1, a counter kept track, dynamically, for them, onscreen, of how many had already been chosen in the 'study' and 'don't study' categories. Data were analyzed, as before, only up until the point that one of these two categories was filled, since once that happened the participants no longer had a choice.

The answer to the question of whether people selected for study the items to which they had not been exposed (and that they therefore had not learned) was positive. As is illustrated in the left panel of Fig. 2, people were much more likely to choose the unpresented pairs than the pairs presented once or six times, $F(2,23) = 34.72$, $p < .0001$, $MSE = .09$, Effect Size = .60. Tukey–Kramer post hoc tests (which were used for all post-hocs throughout this article) showed significant differences between 0 and 1 presentation, and 0 and 6 presentations, though not between 1 and 6 presentations. This case, which we had intentionally made as clear cut as possible, indicated that people declined study of the known items under conditions of new learning.

Because the materials were so easy to learn in a single brief exposure, we had expected little or no difference between the 6- and the 1-presentation conditions. However, in situations in which it would take more time and effort for learning to occur, we expected to find a difference between these two conditions. In the experiment that follows, we used more difficult materials, expecting to find a difference in choice as a function of the degree of learning in these two conditions.

Experiment 3

In this experiment, the only change from Experiment 2 was that the materials we used were unrelated pairs of items such as "footwear-acrobat." These pairs were more difficult to learn than the very high associates of the previous experiment. Participants were 24 Columbia University students who participated for course credit or for pay.

As the center panel of Fig. 2 illustrates, participants chose to study the unpresented items more than the items presented once, which were chosen more than the items presented six times, $F(2,23) = 32.46$, $p < .0001$, $MSE = .09$, Effect Size = .59. Consistent with the increased difficulty in learning, post-hoc tests showed that the 1-presentation condition mean was significantly different from that in the 6-presentation condition.

Experiment 4

Although the results of Experiment 3 showed that people chose to study the once-presented items more than the items presented six times, the difference was not large. In this experiment, we attempted to increase that choice difference, by further decreasing the degree of learning in the 1-presentation condition—making it more like the unlearned 0-presentation condition. To do so, we made the presentation time in the 1-presentation condition shorter, decreasing the time from 2.5 to 1 s. Pairs in the 6-presentation condition were still presented for 2.5 s on each presentation. In all other ways the procedure and materials were the same as Experiment 3. Twenty-four Columbia University students participated for course credit or pay.

As the right panel of Fig. 2 shows, participants chose to study the unpresented items most, followed by items presented once for 1 s and items presented six times for 2.5 s, $F(2,23) = 192.48$, $p < .0001$, $MSE = .02$, Effect Size = .89, and post hoc tests showed that all three groups were significantly different from one another.

We compared the data from Experiments 3 and 4, treating Experiment as if it were a between-participants' factor crossed with the number of study opportunities (1 or 6). There was, of course, a significant effect of number of study opportunities, $F(2,92) = 122.06$, $p < .0001$, $MSE = .06$, Effect Size = .73. More importantly for our hypothesis, the interaction between study opportunities and Experiment was significant, $F(2,92) = 11.62$,

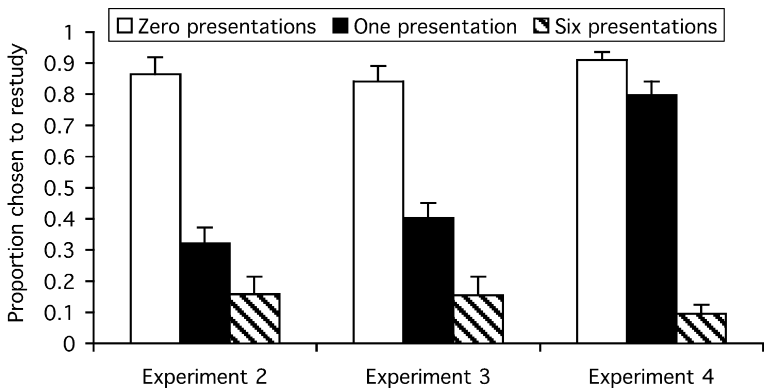

Fig. 2. Proportion of items in the 0-presentation condition, 1-presentation condition, and 6-presentation condition, that were chosen for restudy in Experiments 2, 3, and 4. Experiment 2 used easy-to-learn close associates; Experiment 3 used difficult-to-learn word pairs and allowed 2.5 s for both the 1-presentation condition and the 6-presentation condition. Experiment 4 used difficult-to-learn word pairs, and allowed only 1 s for the 1-presentation condition but 2.5 s for the 6-presentation condition. Error bars indicate standard errors.

$p < .05$, $MSE = .06$, Effect Size = .20. Post hoc t tests were used to further analyze the interaction. They showed that in the 1-presentation condition in Experiment 4 (in which little learning time had been allowed) people were more likely to choose to study, than in the 1-presentation condition, in Experiment 3 (in which more learning time had been allowed), $t(46) = 6.13$, $p < .0001$, while neither 0-presentation condition nor the 6-presentation condition differed between experiments.

These data indicate that people's choices of whether to study follow their learning. When items are easily learned with a single presentation, as in Experiment 2, people decline further study. As learning decreases, as is shown by the contrast of Experiments 3 and 4, peoples choices to study follow. These results all converge on the conclusion that people try to eliminate from further study those items that they already know.

Is the order in which people choose to study unlearned items from easy to difficult or the reverse?

The question addressed here is whether among the items that people believe are as yet unlearned the order of preference is from easy to difficult (as the Region of Proximal Learning model predicts) or from difficult to easy (as the Discrepancy Reduction model) predicts. In the first section, we review studies showing a strong negative correlation between study choice and JOL—seeming to favor the Discrepancy Reduction model. We will argue that these can all be accounted for by the first choice stage of the Region of Proximal Learning model: the correlations are likely to be due to elimination of the already-known items. They do not demand that we infer that among the unlearned items people choose in the order difficult to easy. In the second set of experiments, we look at choice preference that may not be due to simple elimination of the already-known items. In addition, we present two new experiments that focus directly on this issue.

Studies showing a negative correlation between JOL and study choice

Several studies have concluded that, in a choice situation, people prefer the most difficult items for study. When participants are allowed to choose only some items, those that they choose are presumably those to which they are giving preference. For example, Nelson, Dunlosky, Graf, and Narens (1994) in proposing an algorithm that was intended to facilitate learning noted: "The main theoretical assumption contained in the algorithm is that more restudy should be allocated to items that are metacognitively judged to be poorly learned than to items judged to be well learned" (p. 207), as is consistent with the framework of Nelson and Narens (1990). Nelson et al. (1994) reported that the gamma correlation between young adults' JOLs and requested restudy was −.99—indicating that they were, indeed, preferring the poorly learned items. Similarly, Dunlosky and Hertzog (1997) in investigating older and younger adults choice strategies found self-selected gammas were −.80 for their older adults and −.90 for their younger adults. These correlations suggest that they were choosing the lowest JOLs preferentially, though, perhaps, less preferentially for older adults. The question that we raise, in this section, is whether such negative correlations between study choice and JOLs really indicated that people chose to study the items they judged to be most difficult or least well learned, even among the as-yet-unlearned items, or whether these correlations might have come about merely because people eliminated from study those items they already knew that they knew.

In Nelson et al.'s (1994) experiment, participants initially studied pairs of items, made their JOLs, and then were given the opportunity to select half of the items for restudy. Before studying further, though, the participants were tested on the items. They then went on to restudy and be tested on the items over the course of five trials. The results of the first test, made shortly after they made their decisions about whether to restudy or not, showed that participants knew over 45% of the items at the time of the initial test. They should have remembered slightly more of them earlier, at the time when they made their study choices. Requiring the selection of half of the items for restudy, as done in this experiment, is almost exactly the number needed to ensure selection of *everything* that people did not know at the time of judgment, and to not force or even allow discrimination among those unlearned items. If people simply chose all unlearned items, their gamma correlations would, of course, be very high. The authors concluded that "Restudy is allocated more to items that people judge to be poorly learned than to items they judge to be well learned" (p. 212). While the data are consistent with this view, they are also consistent with the view that people simply eliminated those items they believed they already knew.

An experiment comparing younger and older adults using the same procedure as Nelson et al. (1994) was conducted by Dunlosky and Hertzog (1997). These authors also suggested that during choice people selectively preferred the least well-known items. They found that performance on the initial test, which is a conservative estimate of how many items were known during the study choice, was just under 60% for the younger adults, and just under 30% for the older adults. They obtained negative correlations between study choice and JOLs, as had Nelson et al. (1994). It is notable, however, that the magnitude of the negative correlations was smaller for the older adults (who had fewer known items to eliminate) than for the younger adults, just as predicted by the Region of Proximal Learning model. The authors analyzed the probability an item would be chosen given that it was answered incorrectly on the first test: the probability was .84 for younger and .80 for the older adults, so it seems clear that people in both groups were trying to choose for study items they did not know. But whether they preferred the items with the lowest JOL, among these unknown items, cannot be inferred from the data. The observed negative correlations may have resulted simply because people attempted to eliminate the items they knew.

Thiede (1999) allowed people multiple study trials and assigned them the goal of learning everything, with the choice to study as many items as they wanted. On an initial test they answered an average of 31% correct. During the study choice that followed, they selected 70% of the items for restudy. The correlations between study choice and JOL, as in the previously cited studies, were strongly negative (the mean gamma correlation on the first trial was −.74). It is not safe to assume that people selected the most difficult items from among those that were as yet unlearned. Perhaps they only eliminated those items they thought they already knew. In summary, among the choice studies we have been able to find, even those showing very large negative gamma correlations, none have demonstrated that people preferentially chose the least-learned items among those that were not already learned.

Studies favoring the early choice of easy items

Thiede and Dunlosky (1999) conducted a series of experiments investigating the effects of goals on study decisions. They found that when given easy goals, people often chose to study easier items. These experiments are difficult to interpret, however, especially since in some of them, people were actually penalized for remembering too many items. Given that people might have been trying *not* to learn, that they might have been withholding what they knew, or they might have been trying to choose in a way that would be suboptimal in order to thwart their own performance, we can not unambivalently claim these result as favoring the Region of Proximal Learning model.

In Experiment 2, however, Thiede and Dunlosky (1999) asked people to do as well as they could and varied study time. They found that people had a greater tendency to choose the easier items when study time was short, rather than long. With a short total study time people may only have enough time to study some, rather than all, of the items. The short time condition, therefore, provides an indication of which items people choose to study first. The data of Thiede and Dunlosky (1999) support the order choice given by the Region of Proximal Learning model.

Consistent, too, with the results of Thiede and Dunlosky (1999), are data given by Son and Metcalfe (2000). They showed that in several situations, particularly those in which time was limited, people chose to study judged-easy materials earlier than judged-difficult materials. This occurred with long text passages and with sonnets. It is notable that participants in Son and Metcalfe's study were also asked to assess interestingness of each of the passages. They judged the passages they chose to be more interesting than those to which they gave higher judgments of difficulty but did not choose. This finding relates well to the idea, discussed earlier, that people find materials within their Region of Proximal Learning to be engaging.

Metcalfe and Kornell (2003) (and see Metcalfe, 2002, for other cases in which the easiest items are given priority) conducted several experiments in which Spanish–English translations were presented for study three at a time. Within each triad of pairs, one pair was easy (e.g., family–familia), one was of medium difficulty (e.g., turn–volver) and one was difficult (e.g., skylight–buhardilla), and the participants were instructed in advance about where the easy, medium and difficult items would appear on the screen. Either 5, 15 or 60 s were allowed for study of each triad consisting of one easy, one medium, and one difficult pair—forcing a choice among the three types of items. People tended to select the easiest items first, followed by the medium, and then the difficult items: .70 of first choices were easy, .86 of second choices were of medium difficulty, and .69 of third choices were difficult.

Experiment 5

We conducted a follow-up to the Metcalfe and Kornell (2003) experiment in which a similar procedure was used, but in which nine items were presented in a 3×3 grid for 45 s, rather than only three items appearing at one time. In this new experiment, in the leftmost column the three items were easy; in the middle column the three items were of medium difficulty; in the rightmost column the items were difficult. This arrangement allowed us to determine whether or not people were following a reading order, from left to right, repeatedly, as could have been the case in Metcalfe and Kornell (2003), or whether they really preferred the easier items. In Experiment 5, participants were given a pretest, in which all of the English cue words were given on the computer on a single page, and they were asked to type in the Spanish translation for any that they knew. Then they began the choice experiment on the computer. We told participants that "your job is to learn the Spanish translations." As in the previously published experiments, people were shown only the cue items (the English words, in each of the nine positions) and were asked to click, with a mouse, on the item for which they wanted to study the Spanish translation. The computer kept track of the order of clicks and the time on each item. Twenty-eight Columbia University students participated for course credit or pay.

To investigate people's choice order, we coded difficulty as '1' for easy items, '2' for medium-difficulty items,

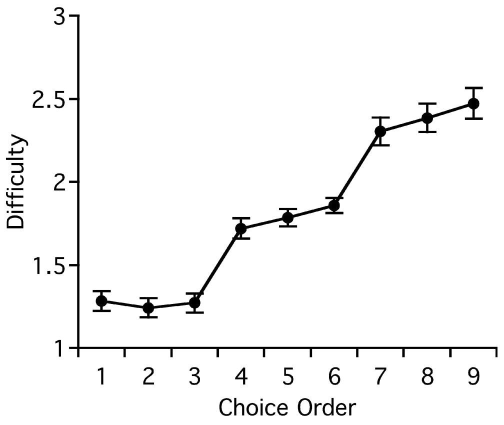

Fig. 3. Level of difficulty of item choices as a function of choice order in Experiment 5. Error bars indicate standard errors.

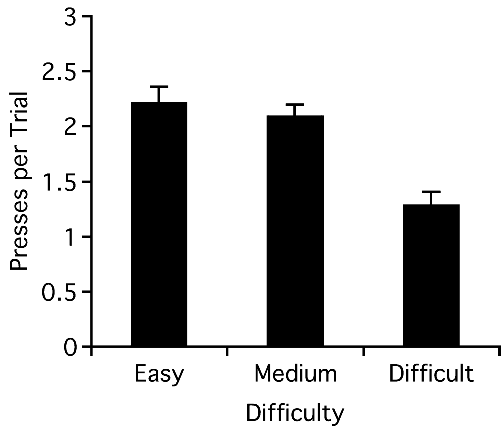

Fig. 4. Mean number of presses per trial for each level of difficulty in Experiment 5. Error bars indicate standard errors.

and '3' for difficult items, and computed the average difficulty of the items selected as a function of choice order. Based on the pretest, items that were known to the participant prior to study were excluded from the analysis. Fig. 3 shows the mean difficulty of the items selected on the first nine choices. A scallop pattern would have resulted if people were selecting from left to right, following the scan pattern used in reading. However, as can be seen from the figure, people tended to select all of the easy items as indicated by mean values near 1, on the first three selections. Then they chose medium-difficulty items (indicated by mean values near 2) on the next three selections. There was a tendency toward higher values on the third three choices, but it is notable that the mean value is considerably less than '3' on these selections—indicating that people did not choose difficult items exclusively, but, rather, often went back to or stayed on the easier items. We conducted an ANOVA on these data treating choice order up to 9 (which allowed us to include all but one participant) as an independent variable. The effect of choice order was significant, $F(8,26) = 90.98$, $p < .0001$, $MSE = .07$, Effect Size = .78. Post hoc tests showed that presses 1, 2, and 3 were not significantly different from each other, nor were 4, 5, 6 or 7, 8, 9. However, there were significant differences between 3 and 4, and 6 and 7. People chose the easiest items first (and stayed on them until they had chosen them all), turning to the more difficult items only later, as predicted.

We also analyzed the total number of times items were selected (Fig. 4). There was a significant effect of difficulty, $F(2,27) = 16.40, p < .0001, MSE = .44$, Effect Size = .38. Post hoc tests showed that easy and medium-difficulty items were selected more often than were the difficult items. In summary, this experiment clearly showed that, when selecting among unknown items, people chose easier items both earlier and more frequently.

Experiment 6

To conclude the section on choice order, we conducted an experiment that was conceptually similar to the experiments of Nelson et al. (1994) and Dunlosky and Hertzog (1997), and to Experiments 1 through 4, above. The main change was that instead of allowing people the option of choosing from among all of the items, including those that they already knew, we forced them to choose among only the *as-yet-unlearned* items.

Each participant received four lists each consisting of 24 English–Spanish translations, which we knew from previous research (Metcalfe, 2002; Metcalfe & Kornell, 2003) varied in difficulty. Eight pairs in each list were easy; eight were of medium difficulty; and, eight were difficult. We first presented the first list of English–Spanish translations, one at a time on the computer screen, for study for 4 s for each pair. Participants were simply told that they "should try to learn the translations, because you will be tested on them later." Immediately following study of each pair, people made a JOL concerning that pair. Once the entire list had been presented and judged, we had the computer administer an immediate test, and score whether people were correct or incorrect on each item.

For the restudy choice phase, people were provided with only the items on which they had given either no answer or an incorrect answer on the original test. These were given, one at a time, in a random order, on the computer screen, and participants were informed that they would be given only items that they had gotten wrong on the test for the choice phase. Participants were allowed to choose half of these items for restudy, and a counter on the screen showed how many items had been chosen, dynamically, in the 'study' and 'don't study' category. If there was an odd number of errors the computer randomly eliminated one item. The Discrepancy Reduction model, of course, predicted that participants should select the most difficult items (i.e., those given the lowest JOLs), which are furthest from being learned. The Region of Proximal Learning model predicted that people should chose those items with the highest JOLs.

The data favored the Region of Proximal Learning model. The items selected for study were judged to be better learned (Mean JOL = 36.70) than those not selected for study (Mean JOL = 25.31), and the difference was significant ($t(47) = 4.24, p < .001$). The Gamma correlation between study choice and JOL (mean = .34) was significantly positive, $t(45) = 4.58$, $p < .0001$. In summary, when selecting among unknown items for study with the goal of maximizing learning, the items closer to being learned were given priority over items further from being learned.

The conclusion from these experiments is that the negative correlations relating choice to JOLs, that have repeatedly been reported, are due exclusively to the first step of the item selection process: people attempt to eliminate from further study the items that they know already. Once those items are eliminated, people proceed in an order from easiest to most difficult, rather than the reverse.

Perseverance

This final section of the paper addresses the issue of how long people persist in studying. The Discrepancy Reduction model's answer to that question, as given by Dunlosky and Thiede (1998), is:

> If the perceived degree of learning has not reached the norm of study, more study time will be allocated to the item. Put differently, this discrepancy-reduction model of self-paced study is based on a negative feedback loop in which study of an item is stopped when the error between the perceived state of learning and the amount of learning desired reaches zero (p. 38).

The data rallied in support of this idea come from the many studies that have shown that people study the items that are more difficult, or the items that they judge to be more difficult, for longer. There is a negative correlation between difficulty or perceived difficulty and study time (Belmont & Butterfield, 1971; Dufresne & Kobasigawa, 1989; Dunlosky & Connor, 1997; Kellas & Butterfield, 1971; Kobasigawa & Metcalf-Haggert, 1993; Le Ny et al., 1972; Mazzoni & Cornoldi, 1993; Mazzoni, Cornoldi, & Marchitelli, 1990; Mazzoni, Cornoldi, Tomat, & Vecchi, 1997; Nelson & Leonesio, 1988; Pelegrina, Bajo, & Justicia, 2000; Thiede & Dunlosky, 1999; Zacks, 1969).

The Region of Proximal Learning model also predicts that there should usually be an increase in study time with difficulty, but for different reasons than does the Discrepancy Reduction model. The amount of time spent studying an item should depend upon when the jROLs approach zero, that is when the perceived information uptake functions level out. Metcalfe and Kornell (2003) provided empirical information uptake functions, over time, for easy- and medium-difficulty items, which should give some idea of when this should happen for various item types.

In Fig. 5, we illustrate idealized information uptake functions for easy, medium, and difficult items. The easy items enjoy rapid information uptake at first but very soon are learned, at which point their jROLs go to zero. When the jROLs are zero, people stop studying. Hence study time will be short with easy items. The uptake functions for medium difficulty items increase slowly and steadily and for a long time, with the jROLs remaining moderate for a long time. Hence study time will be long. The uptake functions for very difficult items are shallow. Whether a person continues to study with such shallow uptake functions will depend on their stopping criterion for the jROLs. Some people may persist for quite a long time if their criterion for stopping is that the jROL be very near zero. Others may quit fairly soon,

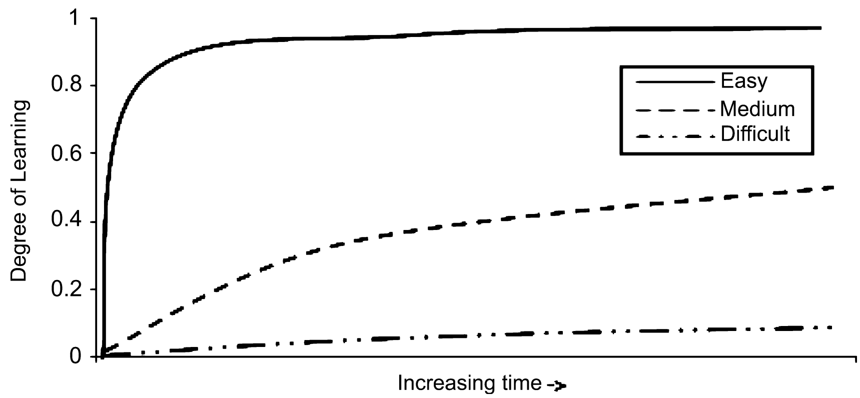

Fig. 5. Idealized information uptake functions for easy, medium-difficulty, and difficult items.

if the stopping jROL value is not as low as the tiny increase shown by the difficult items. Factors such as time pressure, and the difficulty of the other items may also have a particular impact upon whether or not the person persists in studying the very difficult items.

Note that, unlike the Discrepancy Reduction model, the Region of Proximal Learning model does not predict that the study time functions should mirror the amount of time really needed to master the items in each difficulty class, that is, they do not reflect the amount of time required for full learning (see Mazzoni & Cornoldi, 1993; Nelson & Leonesio, 1988), which could be very long indeed in the case of difficult items. The model could accommodate decelerating or even non-monotonic functions relating JOLs to study time. To our knowledge, only one study (Mazzoni et al., 1990) has reported non-monotonic effects (though Mazzoni & Cornoldi, 1993, reported a number of null effects) but this may be because usually data are reported as overall gamma correlations, not segmented by JOL level.

Experiment 7

Despite the many studies showing negative correlations between study time and JOLs few plot out the functions relating the two. We, therefore, conducted Experiment 7. Although we predicted a negative relationship between study time and JOLs, we also expected that there might be a leveling off in the function relating JOLs to study time when the items were extremely difficult because the rate of learning on these items might not be sufficient to exceed the stopping jROL value. We also expected that the stopping jROL would vary across participants. Those who were highly motivated, would have a near zero stopping criterion for their jROLs, and would spend more time overall. Those who were less motivated would have higher stopping jROLs and would spend less time, overall. More importantly than the overall differences in time spent, however, we predicted a selective difference in their perseverance on the judged *difficult* items, since that is where the difference in jROLs would especially come into play.

Our participants were 42 Columbia University students who received course credit or pay for participating. Participants were told that "In this experiment you're going to learn some Spanish vocabulary." We asked them to first study, one at a time, each English–Spanish pair in an 18-item list, consisting of 6 easy, 6 medium-difficulty, and 6 difficult pairs (see, Metcalfe, 2002), for 4 s each, making immediate JOLs after studying each pair. Then there were 3 study/test trials. During these, participants studied each pair, presented in a random order one at a time, for as long as they wanted, then pressed a button when they wanted to go on to the next pair. At the end of each trial, there was a test

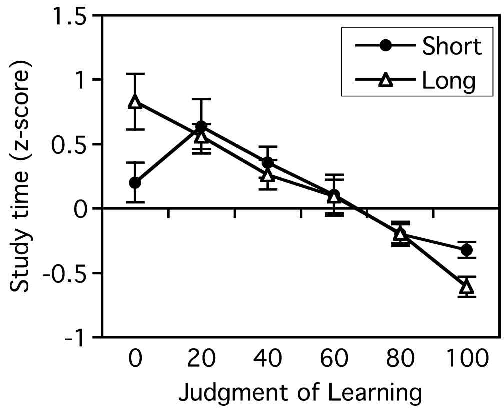

Fig. 6. Study time z scores as a function of immediate JOLs, for participants with long and short overall study times, in Experiment 7. Error bars indicate standard errors.

in which the cues were presented and the participant was asked to type in the correct response. At the end of each test, the computer removed all items that had been correctly answered from the set that would be represented on the next self-paced-study and test trial. There were four lists per participant, each replicating the entire design.

The data were median split separating people who were had long and short overall study times (which, we assumed related to a difference in stopping jROL value) and who had observations in all of the JOL categories (25 participants).[2] To facilitate comparison of the participants with long and short study times, we computed the study time z scores for each participant based on his or her own mean study time, and used these as the dependent measure. Fig. 6 shows the first trial study-time z scores as a function of JOL. As can be seen from the figure, there was an effect of JOL on study time, $F(5,23) = 19.18$, $p < .0001$, $MSE = .22$, Effect Size = .45, as expected from the previous experiments. More pertinent to our hypothesis: People who had long overall times showed a difference from people with short overall times, especially on how long they spent on the very low JOL items, as shown by the significant interaction, $F(5,23) = 2.72$, $p < .05$, $MSE = .22$, Effect Size = .11. The gamma correlations between study time and JOL were $-.34$ and $-.50$ for people with short and long study times, respectively. While both gammas were signifi-

[2] In Experiment 7, we also accumulated all data weighting every observation equally—a method that has the advantage of including all observations and not forcing us to eliminate participants. The z score values (standard errors are in parentheses) split into groups with short and long total study times, for JOLs 0, 20, 40,60, 80, and 100, were: Short: .10 (.12), .34 (.07), .34 (.08), .02 (.06), 0.12 (.07), $-.28$ (.03). Long: .58 (.10), .47 (.07), .20 (.07), .16 (.07), $-.19$ (.05), $-.53$ (.03).

cantly different from zero, $t(19) = -7.01$, $p < .0001$, and $t(20) = -11.49$, $p < .0001$, the gammas were also significantly different from one another, $t(39) = 2.42$, $p < .05$, in the direction predicted by the model.

As these data indicate, our overall findings were consistent with the remainder of the literature insofar as people spent longer on items to which they assigned lower JOLs. However, what these data indicate, that was not apparent from the earlier literature is that study time tapers off with low JOLs, and some people—presumably those with a higher jROL value as a stop rule—study the low JOL items less than the medium JOL items.

Experiment 8

Although much emphasis had been placed on the relation of study time to JOLs made prior to study, in fact, this relation is only tangentially relevant to the stop rule that people use. While it is true that people would need to study the low JOL items longer than the high JOL items to reach a criterion of learning them, as the Discrepancy Reduction model asserts, that does not mean that because people do study the low JOL items longer that they do, in fact, reach that criterion. Indeed, there is every reason to suppose that they do not—their performance on the difficult items is typically far from perfect. Such a result might, of course, be due to poor metacognitions: people think they have learned the items by the end of study, but in fact they have not. But perhaps the reason for the seemingly aberrant data is, instead, that people use a different rule entirely to determine when to stop. To address this issue directly, we need to know people's JOLs when they stop studying.

Since there were no studies in which people were asked for their JOLs at the time they decided to stop studying, we conducted Experiment 8. To determine that the participants were behaving in a way that was not altered by the requirement of indicating their stopping JOLs, we had 19 participants first perform four lists of the experiment outlined above. Their performance was consistent with that in Experiment 7. Then, we added a fifth list, on which we asked them to give their stop JOLs. On the fifth list, just as before, participants studied and made initial JOLs on all items. There were then three study/test trials in which they determined their own study time on each item. When they finished studying each item on each of the three trials, they indicated their JOL for the item they had just finished studying. Then, at the end of each trial, as before, they were tested on the items, and as before the computer removed those items that had been learned, and included the as-yet-unlearned items on the next trial.

Mean study time varied across the five lists in this study (as had also been the case in Experiment 7), $F(4,72) = 3.529$, $MSE = 19.769$, $p = .01$, Effect Size = .85, but the difference was circumscribed to the first list, on which people spent an average of 11.38 s per item. Their study times per item were 8.49, 7.63, 7.18, and 6.43 s for lists 2, 3, 4, and 5, respectively. Tukey–Kramer post hoc tests showed that list two through five were not different from one another, but list 1 was different from both list 4 and 5. Clearly people were speeding up over the session, especially after list 1, but not significantly more in the list in which they made stop JOLs than in the three lists preceding it. Furthermore, the increase in speed, apparently did not impact upon accuracy. The final test accuracy was not different across lists, $F(4,72) = 1.34$, $MSE = .008$, and varied slightly around .82. From these data we inferred that making the stop JOL, explicitly, did little to change how people were studying or learning.

To address the question of whether people use a criterion of learning as their stop rule, we plotted the proportion of responses, collapsed over all participants, at each JOL level for the stop JOLs, for the three trials, as is shown in Fig. 7. We expected to see virtually all of the stop JOLs be 100—indicating that people believed that they had learned the items. However, the majority of the JOLs were less than 100. There was a small peak at 100 on trial 1 (which probably occurred because people actually learned quite a few— mostly easy— items on that trial). However, the other trials showed flat distributions. Even on trial 1, people most frequently stopped before their JOL reached 100. It did not appear

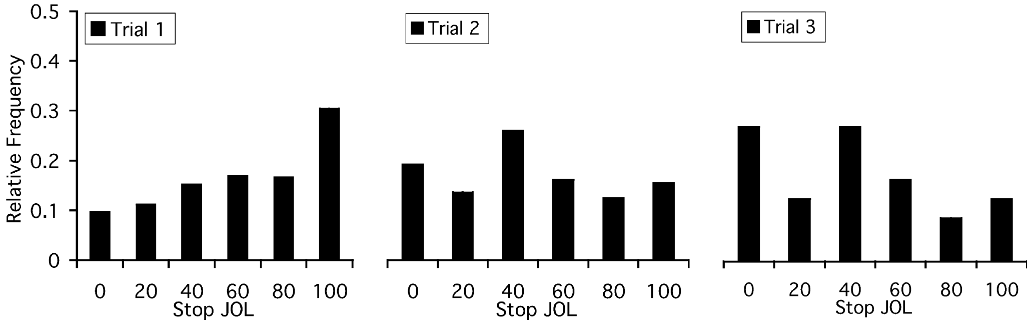

Fig. 7. Relative frequency of Stop JOLs by trial in Experiment 8.

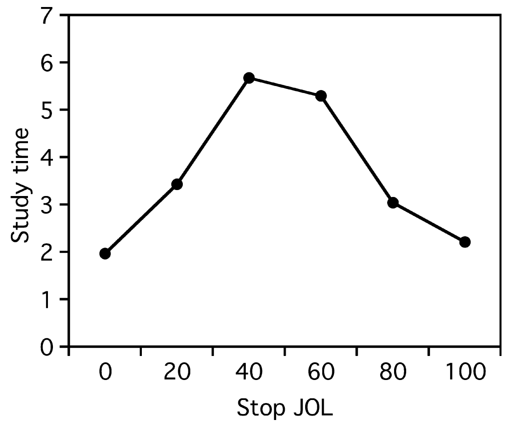

Fig. 8. Study time as a function of Stop JOLs in Experiment 8.

that people's internal criterion was simply lower than 100%, since the JOLs were distributed over the entire range, especially on the second and third trials. These data indicate that people were not using a discernable JOL criterion to stop studying.

We also analyzed the mean stopping JOLs as a function of study time. As Fig. 8 shows, when all observations were weighted equally across participants and lists, people studied for only a short amount of time with very low stopping JOLs as well with very high stopping JOLs. They persisted longer when JOLs at time of stopping were in the midrange.

Finally, to analyze whether people tended to stop when their jROLs were too low we calculated the difference between the initial JOL and the JOL made at the time of stopping. (It would be of interest, of course, to compute jROLs by repeated dynamic sampling of JOLs at short intervals as the person is studying. This would give a more refined measure of the jROL function than merely computing the difference between initial and stop

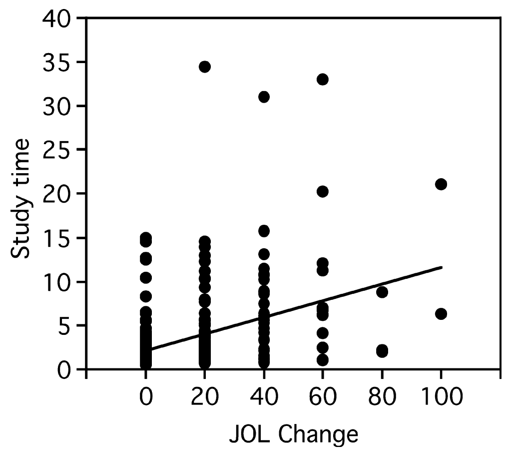

Fig. 9. Study time as a function of the difference between the immediate JOLs and the Stop JOLs in Experiment 8.

JOLs.) Nevertheless, if people continued to study when they perceived themselves to be learning, but stopped when they reached a point of diminishing learning returns, then, despite the crudeness of our measure, we expected to find long study times associated with large changes from initial JOL to stopping JOL. The model makes predictions only about positively valenced differences, so we present them here (but the interested reader can peruse the negatively valenced changes[3]).

The correlation between the change in JOL and study time was computed at the item level (based on all 283 observations of a non-negative change in JOL with each observation from any participant or trial equally weighted). The overall Pearson correlation was .40, $p < .0001$ as is shown in Fig. 9. Longer study times were associated with larger changes in JOLs, consistent with the proposal that the more people think they are learning, the longer they persist.

Conclusion

A strong negative correlation between JOLs and study time has often been thought to indicate that people are using their metacognitions to allocate their study time appropriately. The sometimes rather low values observed have been taken as an indication of poor metacognition or of poor control. In contrast, the Region of Proximal Learning model, presented here, suggests that a strong negative correlation should not be taken as an index of the goodness of metacognitively guided control. One component of the model—that of choosing to not study what is already known—results in a negative correlation. But neither the order of choice once the known items are eliminated, nor the stopping rule indicating when to give up on a particular item, necessitate a negative correlation.

Instead, the Region of Proximal Learning model posits that people first attempt to eliminate items from study the items that are already known. Among the

[3] In Experiment 8, we also repeated the item-level correlation including all data rather than eliminating the data in which the original JOL was higher than the stopping JOL. The correlation was still significantly positive, $r = .31$, $p < .0001$. When we analyzed the negative change data alone, that is the 59 cases in which people's initial JOL was higher than their JOL at time of stopping on the first trial, this correlation still showed a positive, though less strong, trend, $r = .22$, $p = .096$, indicating that the more their JOL changed for the worse, the less they studied. The positive change correlations to which the model applies were also computed separately for each participant, rather than on the item level as above and in the text. As with the item-level analysis, these correlations were also significantly positive (Gamma = .27, $t(18) = 2.55$, $p < .05$, and Pearson $r = .27$, $t(18) = 2.92$, $p < .01$). Neither of the negative change correlations computed by participant, though, were significant (t's < 1).

unknown items, people prioritize from the subjectively easiest to most difficult. Once they are studying an item they must decide when to quit, and turn to something else. The model proposes that the person will continue devoting study time to an item so long as they perceive themselves to be learning, but stop when they feel that learning is no longer paying off.

From the data reviewed and the new experiments presented here, it appears that people behave in a manner that is consistent with the Region of Proximal Learning model. But even if a person behaves exactly as specified by the model, and we would be tempted to say that he or she is exerting good metacognitive control, caution is needed in drawing such a conclusion. The major question left untouched by the present model, which is a model of what people do rather than what they should do, is that of efficacy. We still do not know whether what they do enhances their learning, or is in any way optimal. Until we have answered the still-open question of efficacy, despite the subtlety of people's strategies and their adherence to the predictions of the model, we cannot fully endorse the idea that they are exerting *good* metacognitive control.

References

Atkinson, R. C. (1972). Optimizing the learning of a second-language vocabulary. *Journal of Experimental Psychology, 96*, 124–129.

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

Berlyne, D. E. (1978). Curiosity and learning. *Motivation and Emotion, 2*, 97–175.

Carver, C. S., & Scheier, M. F. (1990). Origins and functions of positive and negative affect: A control-process view. *Psychological Review, 97*, 19–35.

Cull, W. L., & Zechmeister, E. B. (1994). The learning ability paradox in adult metamemory research: Where are the metamemory differences between good and poor learners. *Memory & Cognition, 22*, 249–257.

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

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

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

Dunlosky, J., & Hertzog, C. (1998). Training programs to improve 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., & Nelson, T. O. (1992). Importance of kind of cue for judgments of learning (JOL) and the delayed-JOL effect. *Memory & Cognition, 20*, 374–380.

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.

Hebb, D. O. (1949). *The organization of behavior*. New York: Wiley.

Hertzog, C., Dunlosky, J., Robinson, A. E., & Kidder, D. P. (2003). Encoding fluency is a cue used for judgments about learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 29*, 22–34.

Hyland, M. E. (1988). Motivational control theory: An integrative framework. *Journal of Personality and Social Psychology, 55*, 642–651.

Kellas, G., & Butterfield, E. C. (1971). Effect of response requirement and type of material on acquisition and retention performance in short-term memory. *Journal of Experimental Psychology, 88*, 50–56.

Kimball, D. R., & Metcalfe, J. (2003). Delaying judgments of learning affects memory, not metamemory. *Memory & Cognition, 31*, 918–929.

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

Koriat, A. (1993). How do we know what we know. The accessibility model of feeling of knowing. *Psychological Review, 100*, 609–639.

Koriat, A. (1994). Memory's knowledge of its own knowledge: The accessibility account of the feeling of knowing. In J. Metcalfe & A. P. Shimamura (Eds.), *Metacognition: Knowing about knowing* (pp. 115–135). Cambridge, MA: MIT Press.

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

Koriat, A., & Levy-Sadot, R. (2001). The combined contributions of the cue-familiarity and accessibility heuristics to feelings of knowing. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 27*, 34–53.

Koriat, A., Sheffer, L., & Ma'ayan, H. (2002). Comparing objective and subjective learning curves: Judgments of learning exhibit increased underconfidence with practice. *Journal of Experimental Psychology: General, 131*, 147–162.

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

Lord, R. G., & Hanges, P. J. (1987). A control system model of organizational motivation: Theoretical development and applied implications. *Behavioral Science, 32*, 161–178.

Maki, R. H., & Berry, S. L. (1984). Metacomprehension of text material. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 10*, 663–679.

Masur, E. F., McIntyre, C. W., & Flavell, J. H. (1973). Developmental changes in apportionment of study time among items in a multitrial free recall task. *Journal of Experimental Child Psychology, 15*, 237–246.

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

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

Mazzoni, G., Cornoldi, C., Tomat, L., & Vecchi, T. (1997). Remembering the grocery shopping list: A study on metacognitive biases. *Applied Cognitive Psychology, 11*, 253–267.

Metcalfe, J. (1993a). Novelty monitoring, metacognition, and control in a composite holographic associative recall model: Implications for Korsakoff amnesia. *Psychological Review, 100*, 3–22.

Metcalfe, J. (1993b). Monitoring and gain control in an episodic memory model: Relation to P300 event-related potentials. In A. F. Collins, S. E. Gathercole, M. A. Conway, & P. E. Morris (Eds.), *Theories of memory* (pp. 327–354). Hillsdale, NJ: Lawrence Erlbaum Associates.

Metcalfe, J. (1998). Cognitive optimism: Self deception or memory-based processing heuristics?. *Personality and Social Psychological Review, 2*, 100–110.

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

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

Metcalfe, J., Schwartz, B. L., & Joaquim, S. G. (1993). The cue familiarity heuristic in metacognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 19*, 851–861.

Metcalfe, J., & Shimamura, A. S. (1994). *Metacognition: Knowing about knowing*. Cambridge, MA: MIT Press.

Miner, A. C., & Reder, L. M. (1994). A new look at feeling of knowing: Its metacognitive role in regulating question answering. In J. Metcalfe & A. P. Shimamura (Eds.), *Metacognition: Knowing about knowing* (pp. 47–70). Cambridge, MA: MIT Press.

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, R. J. (1988). Allocation of self-paced study time and the ''labor-in-vain effect''. *Journal of Experimental Psychology: Learning, Memory and Cognition, 14*, 676–686.

Nelson, T. O., & Narens, L. (1980). Norms of 300 general-information questions: accuracy of recall, latency of recall, and feeling-of-knowing ratings. *Journal of Verbal Learning and Verbal Behavior, 19*, 338–368.

Nelson, T. O., & Narens, L. (1990). Metamemory: A theoretical framework and new findings. In G. H. Bower (Ed.). *The psychology of learning and motivation* (Vol. 26, pp. 125–141). New York: Academic Press.

Nelson, T. O., & Narens, L. (1994). Why investigate metacognition. In J. Metcalfe & A. P. Shimamura (Eds.), *Metacognition: Knowing about knowing* (pp. 1–25). Cambridge, MA: MIT Press.

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

Pelegrina, S., Bajo, M. T., & Justicia, F. (2000). Differential allocation of study time: Incomplete compensation for the difficulty of the materials. *Memory, 8*, 377–392.

Piaget, J. (1952). *The origins of intelligence in children*. New York: International Universities Press.

Pirolli, P., & Card, S. (1999). Information foraging. *Psychological Review, 106*, 643–675.

Powers, W. T. (1973). *Behavior: The control of perception*. Hawthorne, NY: Aldine deGruyter.

Reder, L. M. (1988). Strategic control of retrieval strategies. In H. Bower (Ed.). *The psychology of learning and motivation: Advances in research and theory* (Vol. 22, pp. 227–259). San Diego, CA: Academic Press.

Reder, L. M., & Ritter, F. E. (1992). What determines initial feeling of knowing. Familiarity with question terms, not with the answer. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 18*, 435–451.

Rumelhart, D., & McClelland, J. (1986). *Parallel distributed processing: Explorations in the microstructure of cognition*. Cambridge, MA: MIT Press/Bradford Books.

Schwartz, B. L., & Metcalfe, J. (1992). Cue familiarity but not target retrievability enhances feeling-of-knowing judgments. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 18*, 1074–1083.

Schwartz, B. L., & Metcalfe, J. (1994). Methodological problems and pitfalls in the study of human metacognition. In J. Metcalfe & A. P. Shimamura (Eds.), *Metacognition: Knowing about knowing* (pp. 93–114). Cambridge, MA: MIT Press.

Schwartz, B. L., & Smith, S. (1997). The retrieval of related information influences tip-of-the-tongue states. *Journal of Memory and Language, 36*, 68–86.

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

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

Spellman, B. A., & Bjork, R. A. (1992). When predictions create reality: Judgments of learning may alter what they are intended to assess. *Psychological Science, 3*, 315–316.

Stephens, D. W., & Krebs, J. R. (1986). *Foraging theory*. Princeton, NJ: Princeton University Press.

Thiede, K. W. (1999). The importance of accurate monitoring and effective self-regulation during multitrial learning. *Psychonomic Bulletin & Review, 6*, 662–667.

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-regulated study: An analysis of selection of items for study and self-paced study time. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 25*, 1024–1037.

Vygotsky, L. S. (1987). The collected works of L.S. Vygotsky. In R. W. Rieber & A. S. Carton (Eds.). *Problems of general psychology* (Vol. 1). New York: Plenum Press.

Weber, E. U., Shafir, S., & Blais, A.-R. (2004). Predicting risk sensitivity in humans and lower animals: Risk as variance or coefficient of variation. *Psychological Review, 111*, 430–445.

Zacks, R. T. (1969). Invariance of total learning time under different conditions of practice. *Journal of Experimental Psychology, 82*, 441–447.