ELSEVIER

Acta Psychologica 98 (1998) 37–56

acta psychologica

What makes people study more? An evaluation of factors that affect self-paced study

John Dunlosky [a,*], Keith W. Thiede [b]

[a] Department of Psychology, P.O. Box 26164, University of North Carolina at Greensboro, Greensboro, NC 27402-6164, USA

[b] University of Illinois at Chicago, Educational Psychology (M/C 147), 1040 W. Harrison St., Chicago, IL 60607, USA

Received 11 March 1997; received in revised form 28 September 1997; accepted 6 November 1997

Abstract

Allocation of study time across items was investigated in three experiments. According to the norm-affects-allocation hypothesis, when studying an item, a person changes the sought-after degree of learning for the item (called the norm of study) in an attempt to achieve task goals. As the norm of study is increased, more time will be allocated for study. This hypothesis was evaluated by having people pace their study of items for an eventual test of recall. As predicted, study time was greater (a) when points awarded for recalling an item increased, (b) when instructions emphasized mastering each item rather than quickly learning each item, (c) when points deducted for each second of study decreased, and (d) when the likelihood an item would be on the test increased. Also, although allocating more study time was usually accompanied by an increase in eventual recall, under several conditions people's allocation of study time appeared sub-optimal. Discussion focuses on current theory of self-paced study and people's apparent sub-optimal allocation of study time.

PsycINFO classification: 2343

Keywords: Metacognition; Learning

* Corresponding author. Tel.: +1 910 334 5013; fax: +1 910 334 5066; e-mail: j_dunlosky@goodall.uncg.edu.

0001-6918/98/$19.00
PII S0001-6918(97)00051-6

1. Introduction

Self-directed learning may include selecting a particular study activity, monitoring the degree to which new material has been learned, and allocating study time across the to-be-learned material. Perhaps because self-directed learning is central both to everyday memory performance and to achievement in education (Bjork, 1994; Hertzog and Dunlosky, 1996; Schunk and Zimmerman, 1994), numerous aspects of self-directed learning are being actively investigated (for books covering these topics see Hacker et al., in press; Metcalfe and Shimamura, 1994; Nelson, 1992). The present research focuses on an aspect of self-directed learning that may be critical for achievement on memory tests: "Why does a person terminate the study of one item to move on to another?"

One answer to this question is provided by a system model that originates from the test-operate-test-exit unit of Miller et al. (1960) (for descriptions of two frameworks that are the bases of the present discussion see Le Ny et al., 1972, and Nelson and Narens, 1990). A key feature of this model is that a person's allocation of study time is partly controlled by the interplay between two components of metacognition: (a) a person's assessment about the current state of learning for an item and (b) the person's desired degree of learning for the item, which is called a *norm of study*.

When an item is presented for study, a person assesses the degree to which it has been learned and compares this assessment to his or her own norm of study. If the person's perceived degree of learning equals or exceeds the norm of study, the person will terminate study and proceed to the next item. However, 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 (for discussion of system models of control see Powers, 1973).

Because monitoring plays a central role in this model, predictions have often involved relations between people's judgments of learning, which tap perceived degree of learning, and subsequent self-paced study time. As predicted from the *monitoring-affects-control* hypothesis (from Nelson and Leonesio, 1988, which also can be derived from this model), a person will allocate less study time to items judged as more-well learned than to those judged as less-well learned. To test this prediction, participants may study individual items (presented at a fixed rate) and then judge how well each item had been learned. During the next trial, participants pace the study of each item. According to the hypothesis, participant's judgments of learning will be inversely related to subsequent self-paced study time. This outcome has been obtained under a wide variety of conditions (Dunlosky and Connor, 1997; Graf and Payne, 1992; Mazzoni and Cornoldi, 1993; Mazzoni et al., 1990), which provides evidence for the claim that people's monitoring of learning is used to control study time.

In contrast to research investigating whether monitoring is used to control study time, little has been done to evaluate directly the degree to which a person's norm of study influences self-paced study. According to the model, a person's placement of his or her own norm of study critically influences self-paced study time. In particular, if an individual sets the norm of study low for one item and high for another, the former will be studied less. Three experiments were conducted to empirically evaluate the following *norm-affects-allocation* hypothesis: A person will change his or her own norm of study in an attempt to increase the likelihood of achieving the goals produced by the demands of a memory task. As the person increases the norm of study to attain these goals, he or she will allocate more time to studying items. Accordingly, a major goal of the present experiments was to investigate the degree to which people adjust study time (if at all) in response to various task demands. For instance, if the goal is to gain as many points as possible on the upcoming test of memory, the person presumably will set a higher norm of study for items that are worth more points on the test than for items that are worth fewer points (cf. Le Ny et al., 1972). Thus, a prediction is that study time will be greater for items that are worth more points.

The research strategy used here was to provide converging evidence demonstrating the possible necessity of including this decision criterion (norm of study) in theory of self-paced study. The strategy was accomplished by manipulating a variety of factors that a priori would be expected to affect a person's placement of his or her own norm of study, and hence they were predicted to affect study time. Before discussing these factors and specifics of the present experiments, we discuss previous research to provide evidence as to whether people dynamically change their norm of study when allocating study time across individual items.

1.1. Evidence relevant to the norm-affects-allocation hypothesis

Le Ny et al. (1972) speculated that "the process of studying a particular item is dependent upon the motivation state induced through instructions; these determine a 'norm of study' to be reached" (p. 281). This possibility can be evaluated by results from Nelson and Leonesio (1988), who had two groups of students pace their study of individually presented items. One group was instructed to study a given item until they were absolutely sure it had been mastered (called "accuracy-emphasized instructions"), and the other group was instructed to spend only as much time as was needed to learn each item ("speed-emphasized instructions"). According to the norm-affects-allocation hypothesis, people place the norm of study higher after accuracy-emphasized instructions than after speed-emphasized instructions, which in turn will affect self-paced study time. Consistent with this prediction, mean study time was at least twice as great after accuracy-emphasized instructions than after speed-emphasized instructions (Nelson and Leonesio, 1988).

Mazzoni and Cornoldi (1993), Experiment 4 investigated whether study time was affected by anticipating different kinds of test. Participants either studied items in anticipation of a test of recognition or in anticipation of a test of recall. According to the norm-affects-allocation hypothesis, people will set a lower norm of study when studying for a recognition test than for a recall test, with the former yielding less study time. While summarizing their research, Nelson (1993) provided an explanation for the results of Mazzoni and Cornoldi (1993) in terms of a person's placement of a norm of study: "Their Experiment 4 showed that the amount of self-paced study time to items is lower when people are anticipating an upcoming recognition test instead of an upcoming recall test, thereby illustrating how people *change their sought-after degree of learning* as a function of the anticipated retention test" (p. 269, italics added).

Although this effect of anticipating different kinds of test confirms the norm-affects-allocation hypothesis, it may instead be attributed to a person's initial assessment of mastery for an item. In particular, when an item is presented for a recognition test (vs. a recall test), the individual may assess that he or she is already closer to the degree of learning desired. This may occur because people believe that recognition tests are easier than are recall tests (Speer and Flavell, 1979), and hence not as much study time will be needed to achieve the same degree of performance. Thiede (1996) reported evidence consistent with this alternative to the norm-affects-allocation hypothesis. College students studied items in anticipation of either a recognition test or a recall test and made a judgment of learning for each item, which is a person's prediction about the likelihood of correctly retrieving the item on the eventual test. The magnitude of the judgments of learning was greater when students had anticipated recognition than when they had anticipated recall, suggesting that the kind of test anticipated may directly affect people's assessments of learning, which would then be translated into differential study time. Put differently, differential allocation of study across items may be sufficiently described here by a function of an individual's initial assessment rather than by one including both an initial assessment and a dynamically placed norm of study.

Le Ny et al. (1972) examined how awarding points for correctly recalling items would affect study time. During self-paced study, a given item was identified as being worth either 10 points or 1 point, with the notion being that awarding more points will increase a person's norm of study and hence result in longer study times. By manipulating award, research of Le Ny et al. (1972) sidestepped the aforementioned problem based on confounding a person's assessment of learning with the placement of the norm of study (because individuals are unlikely to expect that learning will be more difficult for 10-point items than 1-point items). Results from this research disconfirmed the norm-affects-allocation hypothesis: Study time was not reliably different for items that were awarded 10 points than for items that were awarded 1 point.

In sum, results from Nelson and Leonesio (1988) of self-paced study time confirmed the norm-affects-allocation hypothesis. The results from Mazzoni and Cornoldi (1993) also confirmed this hypothesis; however, a plausible alternative can also account for them. Finally, results from Le Ny et al. (1972) (which sidesteps the aforementioned alternative) disconfirmed the norm-affects-allocation hypothesis.

2. Overview of the present experiments

In the light of this somewhat equivocal evidence, we investigated the effect of four factors on the allocation of self-paced study to further evaluate the norm-affects-allocation hypothesis: (a) instructions that emphasized speed vs. accuracy (Experiment 1, as in Nelson and Leonesio, 1988), (b) the number of points awarded for later correctly recalling an item (Experiments 1 and 2, as in Le Ny et al., 1972), (c) the number of points deducted for each second of time used to study an item (Experiments 2 and 3), and (d) the percentage chance that a given item would later be tested (Experiment 3). Based on the norm-affects-allocation hypothesis, each of these factors was expected to affect people's norm of study and in turn affect study time. Details concerning each factor are discussed immediately prior to the corresponding experiment.

2.1. Experiment 1

One reason that Le Ny et al. (1972) may not have obtained the predicted effect of points on study time is that they "strongly emphasized that [the participant] had to study the items until he knew them *perfectly* in order to make *no error at all* on the test" (p. 283, their italics). Such accuracy-emphasized instructions may have been incompatible with attaining differential norms of study across items (as suggested by Le Ny et al., 1972), given that participants set their norm of study at maximum for each item (regardless of its corresponding award) to ensure mastering the entire list. Accordingly, perhaps if a person's norm of study was not functionally on the ceiling, differential award will affect study time.

This possibility was explored here by manipulating the number of points awarded for correctly recalling items under speed instructions in which people may not attempt to master every item. The rationale is based on the hypothesis that emphasizing speed produces a relatively low norm of study, especially in comparison to the norm of study produced by the accuracy-emphasized instructions used by Le Ny et al. (1972). Under speed instructions, people will maximize the likelihood of gaining the most points by first ensuring that items awarded a greater (vs. fewer) number of points are recalled. Therefore, study time is predicted to be greater for items awarded 10 points than for those awarded 1 point. Although this outcome may seem somewhat intuitive, its importance lies in the fact that it would contrast the lack of an effect of points reported by Le Ny et al. (1972) and would be consistent with the norm-affects-allocation hypothesis. Besides having an experimental group of participants allocate study time under speed instructions, another group allocated study time under accuracy instructions, which allowed us to evaluate whether speed versus accuracy instructions differentially influenced study time (as in Nelson and Leonesio, 1988) using the present methods.

Finally, to examine whether people's allocation of study time differentiates between conditions more with practice on the task, we had each participant complete a single study-test trial of three different lists. Because investigating the effect of multiple trials on study time was secondary to our goal involving analyses of performance on the first trial, statistical analyses of study time are first presented for the critical first trial, and then exploratory analyses across the multiple trials are presented.

2.2. Method

2.2.1. Design and subjects

Two independent variables were manipulated. Instruction to subjects (either accuracy instructions or speed instructions) was a between-subjects manipulation, and points awarded (either 1 or 10 points) was a within-subject manipulation. Eighty students from the University of Washington individually participated to receive extra course credit and were randomly assigned to each group (40 subjects/group) by order of appearance.

2.2.2. Materials

Items were 48 Swahili–English translation equivalents (e.g., *ardhi* – soil; from Nelson and Dunlosky, 1994). These items were separated into three lists of 16 items each. The lists were approximately equated for difficulty of learning. Macintosh computers displayed instructions and items and recorded all responses.

2.2.3. Procedure

Prior to study, participants were instructed to try hard to learn the items, because they would receive points for only items that they eventually recalled; they were also instructed that some items would be worth 10 points whereas others would be worth 1 point. Half of the participants had instructions that emphasized the importance of speed during study: "During study, it is very important that you do not waste time studying each pair. Putt differently, you should not use extra time studying pairs; instead, you should spend as little time as possible to learn each pair." By contrast, the other half of the participants read instructions that emphasized the importance of mastery during study: "During study, it is very important that you continue studying a pair until you are absolutely certain that you have mastered it. Put differently, do not go on to the next pair until you are certain that you will later recall the given pair. Mastery of the list is of utmost important." Moreover, immediately prior to the presentation of each list, participants were instructed to "get ready to study pairs" and were either reminded to "spend as little time as possible learning each pair" (speed-emphasized group) or to "spend as much time as needed to master each pair" (accuracy-emphasized group). Participants were also informed that each list contained 16 items and that they would be tested on a given list before they studied the next.

One study-test trial occurred for items of each list, with the trial for a given list being completed before the study-test trial of the next list. For each participant, the order of the lists for the three trials was randomized. Prior to the presentation of a list, an award of 1 point was assigned to half of the items (randomly chosen) and an award of 10 points was assigned to the remaining items, and then the order of all items was randomized for presentation. During the presentation of an item for study, the point award assigned to an item was indicated directly above it (Le Ny et al., 1972).

Participants were instructed to try hard to learn items so as to earn as many points as possible. When a participant pressed the return key to begin a given trial, the first item was presented. The item remained on the screen until the participant again pressed the return key, which resulted in the presentation of the next item. The time between the presentation of an item for study and the following participant-initiated key press was recorded as the study time for that item, and the amount of time that participants were allowed to study each item was not restricted. To help participants keep track of the time allocated to an item, we had the computer beep each second during the presentation of the item.

After the 16 items of a given list had been studied, a distractor task occurred for 30 s, which was followed immediately by paired-associate recall for those 16 items. During recall, each stimulus was individually presented (e.g., *ardhi* -), and participants were instructed to type the corresponding response (i.e., soil). Omissions were not allowed. The order of items for recall trials was randomized anew from study to recall. After the final item was presented for paired-associate recall, the number of points scored across the 16 items was presented. Immediately after the number of points scored was presented, the study trials of the next list began.

2.3. Results

All differences declared as reliable in the present research have $p < 0.05$.

2.3.1. Self-paced study time

For each participant, we computed the median study time (s) across items separately for each point condition (1 point or 10 points). Means across individual's median study time are reported in the top half of Table 1.

Table 1
Experiment 1. self-paced study time (s) and recall performance

Group/Award	Trial			Overall
	1	2	3	
Self-paced study time				
Speed instructions				
1 point	7.1 (0.66)	8.2 (0.82)	7.2 (0.78)	7.3 (0.71)
10 points	7.7 (0.68)	9.8 (1.0)	9.5 (1.2)	8.6 (0.88)
Accuracy instructions				
1 point	19.6 (2.2)	14.4 (1.5)	13.6 (1.6)	15.0 (1.4)
10 points	19.2 (2.4)	15.4 (1.6)	14.5 (1.7)	15.3 (1.4)
Recall performance				
Speed instructions				
1 point	0.41 (0.04)	0.47 (0.05)	0.43 (0.04)	0.44 (0.03)
10 points	0.50 (0.04)	0.53 (0.05)	0.56 (0.04)	0.53 (0.03)
Accuracy instructions				
1 point	0.62 (0.04)	0.55 (0.05)	0.64 (0.05)	0.60 (0.04)
10 points	0.66 (0.04)	0.67 (0.04)	0.66 (0.04)	0.66 (0.03)

Entries are mean across individual's median self-paced study time (s) and the mean across individual's proportion of correct recall performance. Values in parentheses are standard errors of the means.

The effect of points and instructions on study time for the first study-test trial were assessed by a (2) (points awarded) × 2 (instructions) analysis of variance (ANOVA). The main effect of points was not reliable, $F(1,78) = 0.03$, MSE = 7.62. Study time was substantially greater when accuracy was emphasized than when speed was emphasized, $F(1,78) = 26.0$, MSE = 221.9. Finally, the interaction was not reliable, $F(1,78) = 1.50$, $p = 0.23$, even though the effect of points on study time had an opposite trend after the two kinds of instruction.

A (2) × 2 × (3) ANOVA was conducted to assess the effects of points and instructions across the three trials. A main effect was reliable for points, $F(1,78) = 7.10$, MSE = 16.8, and for instructions, $F(1,78) = 22.87$, MSE = 327.2, but was not reliable for trial, $F(2,156) = 2.65$, MSE = 74.0. The Instruction-by-Trial interaction was reliable, $F(2,156) = 6.87$, indicating that study times remained relatively constant across trials for the speed-emphasized group but decreased across trials for the accuracy-emphasized group. The Trial-by-Points interaction was also reliable, $F(2,156) = 4.73$. This interaction indicates that study times became increasingly larger for 10-point items than for 1-point items across the three trials. However, because the effect of trials was consistently small and not relevant to the central goals of the present research, these effects were not investigated further. All other interactions were not reliable, $F\text{s} < 1.80$.

2.3.2. Recall performance

For each participant, the proportion of items that were correctly recalled was calculated separately for each point condition. Means across individual's proportion of correctly recalled items are reported in the bottom half of Table 1. A (2) × 2 × (3) ANOVA was conducted as in the analysis of study time. As apparent from inspection of Table 1, recall performance was reliably greater after accuracy-emphasized instructions than after speed-emphasized instructions, $F(1,78) = 10.67$, MSE = 0.25. Although the effect of points on recall performance appears relatively small, it was reliable, $F(2,156) = 21.57$, MSE = 0.03. The main effect for trial and all interactions were not reliable, $F\text{s} < 2.20$.

2.4. Discussion

Instructions had a substantial and consistent effect on self-paced study, with total study time being almost twice as great for people given accuracy-emphasized instructions than for those given speed-emphasized instructions. Although the interaction between instructions and points was not reliable during the first trial, the overall effect of points awarded on study time is consistent with the notion that people place a norm of study under some conditions. In contrast to these effects, however, awarding 9 extra points for correct recall rarely increased study time by even 2 s. One reason why the effect of points may have been relatively small is that participants perceived the difference in points as inconsequential to maximizing their score (i.e., only a 9-point difference both in the present experiment and in Le Ny et al., 1972). Some participants may have not even attended to the points being awarded, especially because recalling all items correctly would yield the maximum score.

Increases in study time after accuracy-emphasized instructions (vs. speed-emphasized instructions) were accompanied by reliable gains in recall, and increases in study time for 10-point items (vs. 1-point items) were often accompanied by reliable gains in recall. These *labor-and-gain* effects contrast *labor-in-vain* effects of Nelson and Leonesio (1988) in which increases in self-paced study time were accompanied by unreliable increases in recall performance. Even though the present method yielded labor-and-gain effects, consider again that participants in the accuracy group were instructed to master each item. Participants in this group used substantially more study time, but they achieved a level of recall (64% overall) that was far from perfect. Thus, although this group's performance appeared to benefit some from extra study, they failed to achieve their performance goal of mastering every item.

2.5. Experiment 2

In Experiment 2, two factors were manipulated that were expected to affect people's norm of study. First, because points awarded had a relatively small effect on study time in Experiment 1, we attempted to increase the salience of this manipulation by having a wider range of points (8, 16, or 64). Although there is an eightfold increase from the 8-point award to the 64-point award in Experiment 2 and a somewhat larger tenfold increase from the 1-point award to the 10-point award in Experiment 1, the range of points awarded in Experiment 2 is substantially larger than the range of points awarded in Experiment 1. Because people's choices are arguably influenced more by differences in the range of awards than by differences in the ratio of awards (as suggested by research in decision making, e.g., see John et al., 1983; Wright, 1984), the wider range of points used in Experiment 2 will presumably be more salient and hence may yield a larger, more robust effect on self-paced study times.

Second, we manipulated the points deducted for each second of time used during study (called "cost"). Imposing a cost for each second of study time was expected to produce an effect similar to speed-emphasized instructions. That is, as compared to participants who were not penalized for using study time, participants who lost points for each second of study were expected to set a lower norm of study and in turn spend less time studying items. Imposing a cost may also make differences in points more salient to participants. Spending more than a few seconds on 8-point items may often yield a loss in points (assuming recall performance is far from perfect), whereas focusing more study time on 64-point items may more often yield gains in points.

Finally, one group of participants was instructed to use rote repetition, and another was instructed to use interactive imagery. The kind of study activity was manipulated because using rote repetition vs. interactive imagery may moderate the presence/absence of labor-and-gain effects. The idea here is that extra study time will benefit recall performance more when people use a relatively effective learning strategy than when they use a relatively ineffective one (for similar rationale concerning boundary conditions for labor-in-vain effects, see Nelson and Leonesio (1988), p. 685). If so, labor-and-gain effects will occur after interactive-imagery instructions and smaller labor-and-gain effects (and perhaps even labor-in-vain effects) will occur after rote-repetition instructions. For instance, assuming that deducting no points during study (vs. deducting points) will elicit longer study times, recall performance will be greater when no points are deducted (vs. when points are deducted) under imagery instructions but not under repetition instructions. The use of ineffective learning strategies may also partly explain why substantial boosts in study time yielded poor recall in Experiment 1. Without training, people may rarely use mediators to study Foreign-language translation equivalents like those used in Experiment 1. Thus, when no points are deducted for studying 64-point items (i.e., when study times will presumably be the longest), recall performance under imagery instructions may be close to perfect.

2.6. Method

2.6.1. Design and subjects

Points awarded for eventually recalling a given item was a within-subject manipulation with three levels (8,16 or 64 points). Cost (0 or 1 point deducted per second of study) and the instructions for study (interactive imagery vs. rote repetition) were between-subjects manipulations. For half the participants, no cost was associated with study time (as in Experiment 1). For the other half, 1 point was deducted for each second of study.

One-hundred and sixty undergraduate psychology students from the University of Illinois at Chicago participated individually for class credit. Forty participants were randomly assigned to each of the four between-subjects groups (cost by instructions) by order of appearance.

2.6.2. Materials

Because the short lists used in Experiment 1 included relatively few observations per subject that may have yielded somewhat noisy data, we increased the number of items to 75. Each pair was constructed from two unrelated, concrete nouns (e.g., barrel – star) so it would be possible to use interactive imagery. Macintosh computers displayed instructions and items and recorded all responses.

2.6.3. Procedure

Twenty-five items were assigned to each of the three point values (8,16, and 64). Items were then randomly ordered for study with the only constraint being that no more than two consecutive items had the same point value. Prior to study, participants were instructed to try hard to learn items so as to earn as many points as possible. After instructions were presented, the self-paced study trial (with the point award assigned to each item being shown directly above it during study) and the test trial were conducted as in Experiment 1.

2.7. Results

2.7.1. Self-paced study time

For each participant, the median study time (*s*) across items was computed separately for each condition. Means across individual's median study time are reported in Table 2. A (3) × 2 × 2 ANOVA was conducted to examine the effects of points awarded (8 vs. 16 vs. 64 points), cost (0 vs. 1 point deducted per second of study time), and instructions (rote repetition vs. imagery).

Several results are apparent from inspection of Table 2. First, more time was allocated for study when no points were deducted for study time than when a point was deducted for each second of study, $F(1,156) = 13.02$, $MSE = 127.62$. Second, the main effect for study activity was not reliable, $F(1,156) = 0.28$, $MSE = 127.62$.

Finally, as in Experiment 1, the effect of points on study time was relatively small but reliable, $F(2,312) = 16.46$. Follow-up *t*-tests were collapsed across groups because no interactions were reliable, $F\text{s} < 1.40$. Study time was reliably greater for 64-point items than for 16-point items, and it was reliably greater for 16-point items than for 8-point items, $t(159)\text{s} > 2.85$.

2.7.2. Recall performance

For each participant, the proportion of items that were correctly recalled was calculated separately for each point condition. Means across individual's proportion of correctly recalled items are reported in Table 3.

A (3) × 2 × 2 ANOVA was conducted as in the analysis of study time. Recall performance was greater when no points were deducted for study time than when a point was deducted for each second of study, $F(1,156) = 7.19$, $MSE = 0.13$. Recall was less for participants who were instructed to use rote repetition than for participants who were instructed to use interactive imagery, $F(1,156) = 12.72$, $MSE = 0.13$. The effect of points on recall performance was also reliable, $F(2,312) = 13.19$. Follow-up *t*-tests were collapsed across groups (cost and instructions) because no interactions were reliable, $F\text{s} < 1.00$. Recall performance was reliably greater for 64-point items than for 16-point items, and it was reliably greater for 16-point items than for 8-point items, $t(159)\text{s} > 2.75$.

Table 2
Experiment 2. self-paced study time (*s*)

Group/Award	Points deducted per second of study time	
	0 (no cost)	1 (cost)
Rote repetition		
8	10.1	6.5
16	11.4	7.1
64	12.2	7.6
Interactive imagery		
8	10.9	7.6
16	11.3	7.9
64	11.8	8.6

Entries are the mean across individual's median study time (*s*). All SEMs $\leqslant 1.4$. The award was presented during the study of each item and is the number of points that was awarded for an item that was correctly recalled.

Table 3
Experiment 2. eventual recall performance

Group/Award	Points deducted per second of study time	
	0 (no cost)	1 (cost)
Rote repetition		
8	0.29	0.16
16	0.30	0.19
64	0.32	0.23
Interactive imagery		
8	0.39	0.31
16	0.41	0.33
64	0.42	0.35

Entries are mean across individual's proportion correct recall performance. All SEMs $\leqslant 0.04$. The award was presented during the study of each item and is the number of points that was awarded for an item that was correctly recalled.

2.8. Discussion

Although the effect of points on study time confirms the norm-affects-allocation hypothesis, the relatively small effects indicate that individuals may not allocate study time across items in a normatively ideal fashion, especially when points are deducted for each second of study. For instance, under repetition instructions, participants studied 8-point items an average of 6.5 s/item, whereas 64-point items received an average of only 7.6 s/item. Considering the level of recall for these conditions (Table 3), the average participant had a net loss of about 131 points for 8-point items but had a net gain of about 178 points for the 64-point items. [1] Why spend any time studying the 8-point items, and why not spend more time on 64-point items? Answers to this question are considered in Section 3.

Deducting no points per each second of study (vs. deducting 1 point/s study) increased study time and recall performance, again demonstrating labor-and-gain effects. These effects of cost were also evident both after imagery instructions and after rote-repetition instructions. However, even after instructions to use interactive imagery, the magnitude of recall performance never exceeded 50% – even under conditions that elicited the longest study times. In sum, the results confirm the norm-affects-allocation hypothesis, but they also demonstrate what appears to be relatively ineffective allocation of study time.

2.9. Experiment 3

In Experiment 3, another factor was examined that was expected to affect a person's norm of study: the percentage chance a given item would be on the test. Use of

[1] For each condition, the average number of points awarded can be derived with values from the tables and the equation, $rpi - sic$. For a given condition, r is the proportion of correct recall performance, p is the points awarded, i is the number of items, s is the number of seconds allocated to study, and c is the cost in points/s of study.

this manipulation was inspired from educational settings in which students ask about the likelihood that a particular topic will be covered on an upcoming examination. According to the norm-affects-allocation hypothesis, participants will set a higher norm of study for items that are most likely to be tested to increase the likelihood of attaining a high-performance goal. The idea here is that students realize that all items will not be learned during one study trial, and hence they will set a higher norm of study for items that are most likely to contribute to test performance. Based on this rationale, study times were predicted to be greater for items that were most likely to be tested (90% chance of being tested) than for items that had an intermediate chance of being tested (50% chance), which will in turn be studied more than items that were least likely to be tested (10% chance). Finally, because the effects of cost on study time had not been examined prior to the present research, we examined the effects of this factor again in Experiment 3.

2.10. Method

2.10.1. Design, subjects, and materials

Cost (0 or 1 point deducted per second of study) was a between-subjects manipulation as in Experiment 2. Percentage chance of an item being on the test was a within-subject manipulation: A third of the items were described as having a 10% chance of being on the test, another third were described as having a 50% chance of being on the test, and another third were described as having a 90% chance of being on the test.

Eighty undergraduate psychology students from the University of Illinois at Chicago participated individually for class credit. Forty participants were assigned to each of the two groups by order of appearance. Materials were the same as those used in Experiment 2, with Macintosh computers displaying instructions and items and recording all responses.

2.10.2. Procedure

Participants were instructed that 8 points would be awarded for each item that was correctly recalled, and that they should study the items so as to earn as many points as possible. After instructions were presented, participants studied items at their own pace as in the previous experiments. During the study of a given item, the percentage chance of that item being on the test was presented above it. Items were randomly ordered for study with the only constraint being that no more than two consecutive items were assigned with the same percentage chance of being tested.

Following presentation of the last item for study, items were randomized anew and presented for paired-associate recall as in the previous experiments. All of the items were presented on the test to have the same number of observations across conditions. Note, however, that participants were not told that they would be tested on all items, and hence the critical effect of chance on study time would not be influenced by testing all items.

2.11. Results

2.11.1. Self-paced study time

For each participant, the median study time (*s*) across items was computed separately for each condition. Means across individual's median study time are reported in the top half of Table 4.

A (3) (chance) × 2 (cost) ANOVA was conducted. As in Experiment 2, study time was reliably greater for participants who did not have points deducted for each second of study time than for those who did have points deducted, $F(1,78) = 16.2$, $MSE = 75.2$. A main effect occurred for chance, $F(2,156) = 11.5$, $MSE = 3.59$. Follow-up *t*-tests were collapsed across the two cost groups because the Cost-by-Chance interaction was not reliable, $F(2,156) = 1.10$. Study time was reliably greater for 90%-chance items than for 50%-chance items, and it was also reliably greater for 50%-chance items than for 10%-chance items, $t(79)$s > 2.10.

2.11.2. Recall performance

For each participant, the proportion of items that were correctly recalled was calculated separately for each condition. Means across individual's proportion of correctly recalled items are reported in the bottom half of Table 4.

A (3) × 2 ANOVA was conducted as in the analysis of study time. Recall performance was greater when no points were deducted for study time than when a point was deducted for each second of study time, $F(1,78) = 4.01$, $MSE = 0.08$. The effect of chance was also reliable, $F(2,156) = 18.3$, $MSE = 0.01$. Follow-up *t*-tests (collapsed across cost groups) revealed that recall performance was reliably greater for 90%-chance items than for 50%-chance items, and it was also reliably greater for 50%-chance items than for 10%-chance items, $t(79)$s > 3.20. The Cost-by-Chance interaction was not reliable, $F < 1.0$.

Table 4
Experiment 3. self-paced study time (*s*) and eventual recall performance

Chance	Points deducted per second of study time	
	0 (no cost)	1 (cost)
Self-paced study time		
10	8.2	4.2
50	9.6	4.8
90	10.0	5.3
Eventual recall performance		
10	0.22	0.14
50	0.25	0.19
90	0.30	0.22

Entries in the top half of Table 4 are the mean across individual's median study time (*s*), and entries in the bottom half of Table 4 are mean across individual's proportion of correct recall performance. Chance indicates that when a given item had been presented for study, participants were presented this value as the percentage chance of the item being on the test. For study time, all SEMs $\leqslant 1.2$; for eventual recall performance, all SEMs $\leqslant 0.04$.

2.12. Discussion

The reliable effects of cost and chance on study time provide converging evidence for the norm-affects-allocation hypothesis. These two factors also produced labor-and-gain effects. Perhaps surprisingly, when participants were told that items had only a 10% chance of later being tested, they still spent a substantial amount of time studying those items, suggesting that individuals do not always allocate study time in a normatively ideal fashion.

3. General discussion

In contrast to Le Ny et al. (1972), in Experiments 1 and 2 the number of points awarded for correctly recalling an item had a relatively small (but reliable) effect on people's allocation of study time. There was one notable exception, however: When participants were given instructions that emphasized mastering each item, self-paced study on the first study trial was not reliably greater for items awarded 10 points (19.2 s) than for items awarded 1 point (19.6 s). Thus, although results from Experiment 1 were not statistically conclusive (i.e., the Instruction-by-Points interaction for Trial 1 was evident but not statistically reliable), the trend was consistent with the hypothesis that Le Ny et al.'s mastery instructions were incompatible with attaining differential norms of study across items. Namely, on the first trial, participants set their norm of study at maximum for each item (regardless of its corresponding award) to ensure mastery of the list as well as maximal points. This possibility also supports a post-hoc explanation for other outcomes from Experiment 1. During the recall test of trial 1, participants presumably monitor recall performance and realize that mastering a list is unlikely. If so, those who had received accuracy instructions may then lower their norm of study, which will result in both (a) the allocation of less study time on subsequent trials and (b) differential allocation of study to 1-point items vs. 10-point items.

Although further research is needed to more fully evaluate the conclusions above, results from the present experiments rule out one reason why Le Ny et al. (1972) did not find an effect of points on study time. They speculated that the number of points awarded did not affect self-paced study time because money was not awarded in exchange for points. That is, points alone were "not meaningful enough" (Le Ny et al. (1972), p. 287) to affect people's norm of study. Even though exchanging money for points may produce a greater effect of awards on study time than those reported here, the present research demonstrated that awarding money is not necessary to affect people's allocation of study time.

More important for the focus of the present research, effects reported in all three experiments confirm the norm-affects-allocation hypothesis. Points awarded, speed-vs.-accuracy instructions, cost for studying, and the likelihood an item would appear on the test all affected study time as predicted. These diverse findings fit well into the discrepancy-reduction model of self-paced study in which people dynamically place a norm of study for each item. Based on the present findings, this theoretical model also provides a preliminary answer to the question posed in the title of this article. Namely, certain factors increase the degree of learning that a person seeks for a given item, which in turn is translated into the allocation of more study time.

A dynamically placed norm of study may be critical both in theory of studying discrete items as well as in theory of how people allocate time when studying text material. For one example, LaPorte and Nath (1976) had participants study a paragraph for an upcoming test. Some participants were instructed to read the paragraph so that they would get 90% of the test questions correct (called "hard goal"), whereas others were instructed to get 25% of the questions correct (called "easy goal"). This instructional manipulation was crossed with presentation rate. Some participants paced their study, and others had an experimenter-set amount of time for study. For participants who had paced their study, study time (as well as test performance) was greater after hard-goal instructions than after easy-goal instructions, which is consistent with a norm-affects-allocation hypothesis for studying text. However, when the presentation rate was experimenter paced, recall performance was also greater after hard-goal instructions than easy-goal instructions. This outcome suggests that participants given a harder goal studied the paragraph differently from those given an easier goal (LaPorte and Nath, 1976). Thus, perhaps the differential study time shown by the self-paced groups was partly an epiphenomenon of using different study activities. Under hard-goal instructions, people may use more-effective study activities that boost recall performance but also take more time to complete.

Similarly, the factors investigated in the present research may have had a qualitative effect on how people studied the items. If so, changes in study activity across levels of a factor may have mediated both study time and recall performance. For instance, when no points were deducted for study, people may realize they have unlimited study time to learn each item and hence select a more effective study activity (or use multiple activities) even though this may require more time to complete. Thus, differential study time may reflect how cost affected the kind of activity used and not changes in people's norm of study. This hypothesis highlights the significance of findings from Experiment 2. In particular, even when participants were instructed to use the same study activity for every item of a list, both cost and award affected people's allocation of study time as predicted by the norm-affects-allocation hypothesis.

Although the focus of this research was more on self-paced study time than on recall performance, the pattern of findings involving both measures provide insight into the labor-in-vain effect in which increases in self-paced study time produce little, if any, increases in recall performance. Nelson and Leonesio (1988) developed two possible boundary conditions for the labor-in-vain effect. First, it may be a function of the kind of study activity people use during self-paced study. According to this account, using a relatively ineffective study activity was expected to yield labor-in-vain effects. In contrast, labor-and-gain effects occurred (a) for Swahili–English translation equivalents that people may not be able to consistently use elaborative study activities without training and (b) even after instructions to use rote repetition. Second, the labor-in-vain effect may be a function of the stage of learning, with the effect occurring during single study trials and not occurring during multiple trials in which feedback is provided by testing. In all three experiments, labor-and-gain effects were obtained after a single study trial. Thus, although the use of relatively ineffective strategies or a single trial may partly account for labor-in-vain effects, these conditions alone will not necessarily lead to labor in vain.

In contrast to labor-in-vain effects, participants here allocated their study time in ways that potentially boosted performance, such as by using less study time when points were deducted for each second of study. However, another kind of labor-in-vain effect was also evident. In particular, participants allocated too much study time in some situations where pay-offs were minimal. In Experiment 2, when a cost of 1 point/s was deducted during study and 8 points were awarded for eventual recall, individuals spent 6.5 s studying items after rote-repetition instructions and 7.6 s studying items after imagery instructions. Thus, on average, people lost 162.5 points and were awarded only 32 points after repetition instructions, and they lost 190 points and were awarded only 62 points after imagery instructions. They would have gained points if recall had been near perfect, but it never exceeded 50%. In Experiment 3, when points were deducted for studying, people used about 4.2 s to study each item when there had been only a 10% chance of an item appearing on the test. Given that the proportion of correct recall was only 0.14, participants had a net loss of 77 points for these items.

In these cases, using any study time may be considered a kind of labor in vain because the points awarded for recalling a few items will not compensate for the cost of studying along with low levels of recall. Study time may have been more optimally used if participants had quickly moved on to the next item when little, if any, points would have been gained through study (cf. "fast-no" responses in recognition tasks; Kolers and Palef, 1976). That is, the norm of study for these items would be more optimally set at 0. Why do people persist in studying when little to nothing will be gained?

One possibility is that this sub-optimal use of study time is more due to the speed of an initial decision than to inadequately allocating study across items. That is, an individual may take several seconds to decide not to study an item. To capture this component of self-paced study, the present model would require a decision about whether to engage study for a particular item (cf. Barnes et al., in press; Koriat and Goldsmith, 1996; Reder and Ritter, 1992), which would occur prior to the discrepancy-reduction process that presumably occurs while allocating study time to a selected item. Note, however, that a slow initial decision of whether to select an item prior to studying it contrasts how quickly people appear to make other metacognitively based decisions. For instance, Reder and Ritter (1992) estimated that people decide whether to retrieve (versus to compute) an answer to a math problem in less than 1 s. Thus, although the present selection-plus-allocation account of self-paced study is plausible, empirical evaluations of two core assumptions of this account – that an initial decision is made prior to allocation and that the decision requires several seconds – are needed.

Another possibility focuses on the accuracy of people's on-line judgments of learning while studying. When people are asked to make judgments of learning during the study of an item (or immediately after an item had been studied), people's judgments are not highly accurate at predicting the likelihood of subsequent recall performance (Nelson and Dunlosky, 1991). Several hypotheses for this poor predictive accuracy have been offered (for other hypotheses not described here, see Dunlosky and Nelson, 1997). Begg et al. (1989) hypothesized that people base their judgments of learning on the ease with which an item is processed during study. More recently, Mazzoni et al. (1997) argued that people's judgments of memory are partly biased by the use of heuristics, such as by judging one's memory of an item by how representative that item is of a larger class. Because these bases for judgments of learning are often not highly correlated with subsequent memory performance, people's judgments will not be highly accurate.

Given this poor predictive accuracy, perhaps people judged that they were learning items well enough to compensate for the cost of studying them, even in those conditions in which the normative utility of study was close to nil. This inadequate-monitoring hypothesis also provides an explanation for other results, such as the lack of mastery given unlimited time to study and labor-in-vain effects. In both cases, people may inaccurately judge that each item had been well-learned during study and hence prematurely terminate self-paced study (for further discussion, see Nelson and Leonesio, 1988). Evidence from Experiment 1, however, suggests that poor monitoring cannot fully account for all kinds of sub-optimal allocation. Participants studied three different lists during consecutive study-test trials. Performance monitoring during the first test trial would provide highly accurate feedback that the goal of mastery had not been met. Thus, one would expect study times to increase across trials until participants met the goal on a given trial. Put differently, after participants realize that their initial monitoring during study under-estimated subsequent recall, they would attempt to overcome this discrepancy by studying more. However, as shown in Table 1, the opposite pattern occurred. This outcome suggests another mechanism that may contribute to sub-optimal allocation: Lack of mastery given unlimited study time and labor-in-vain effects may partly occur because people believe they will not be able to master the list and hence terminate study prematurely.

Finally, we offer one alternative to the system model of self-paced study that was described in the Introduction. Instead of conceptualizing the stopping rule as a function of one's perception of a specific state of learning, the stopping rule may be a function of perceived change in learning. Termination of study will occur for an item when a person perceives no change in learning for a set amount of time, t. That is, a person will continue studying as long as he or she infers that progress is being made within a set amount of time. The various factors that affected study time in the present experiments would be hypothesized to affect the setting of t. Concerning the sub-optimal allocation of study time (e.g., spending over 6 s on 8-point items in Experiment 2), during the first few seconds of study, people may perceive a relatively fast rate of change in learning, and hence will not terminate study even if t is set short. This model also can account for the negative relation between people's judgments of learning and study time demonstrated in previous research and Nelson and Leonesio's labor-in-vain effect. For the latter, when given instructions to master a list (accuracy instructions), t is set for a long duration, so people will continue on task for a long period of time even after they initially perceive that no progress is being made. Even though this alternative model is speculative, the idea that people monitor changes is not new (e.g., Carver and Scheier, 1990), and it provides a plausible alternative to the kinds of discrepancy-reduction model that dominate theory of metacognitive control.

In summary, findings from the present research contribute to our knowledge about self-paced study by demonstrating both the role of a dynamically placed norm of study in the theory of self-paced study and various conditions that yield sub-optimal allocation. Concerning the latter, people did not attain mastery after unlimited study time and appeared to allocate too much study time to items that would have a minimal expected value of award. Considering the prominent role of self-directed learning in everyday life, determining why such failures occur will have important implications not only for theory but also for many real-world applications.

Acknowledgements

An earlier draft of this manuscript was written while the first author was supported as a post-doctoral fellow by NIA Grant T32 AG00175 from the Georgia Institute of Technology. We thank Bennett Schwartz for helpful comments on a draft of this article.

References

Barnes, A.E., Nelson, T.O., Dunlosky, J., Mazzoni, G., Narens, L., in press. An integrative system of metamemory components involved in retrieval. In: Gopher, D., Koriat, A. (Eds.), Attention and Performance XVII. Cognitive regulation of performance: Interaction of Theory and Application. MIT Press, Cambridge, MA.

Begg, I., Duft, S., Lalonde, P., Melnick, R., Sanvito, J., 1989. Memory predictions are based on ease of processing. Journal of Memory and Language 28, 610–632.

Bjork, R.A., 1994. Memory and metamemory: Considerations in the training of human beings. In: Metcalfe, J., Shimamura, A.P. (Eds.), Metacognition: Knowing About Knowing. MIT Press, Cambridge, MA.

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

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., Nelson, T.O., 1997. Similarity between the cue for judgments of learning (JOL) and the cue for test is not the primary determinant of JOL accuracy. Journal of Memory and Language 36, 34–49.

Graf, E.A., Payne, D.G., 1992. Assessing the scope of the delayed-JOL effect: Is it evident in the prediction of the self-paced allocation of study? Unpublished Honors Thesis, State University of New York at Binghamton.

Hacker, D.J., Dunlosky, J., Graesser, A.C. (Eds.), in press. Metacognition in Educational Theory and Practice. Lawrence Erlbaum, Hillsdale, NJ.

Hertzog, C., Dunlosky, J., 1996. The aging of practical memory: An overview. In: Hermann, D.J., McEvoy, C., Hertzog, C., Hertel, P., Johnson, M.K. (Eds.) Basic and Applied Memory Research: Theory in Context, vol. 1. Lawrence Erlbaum, Hillsdale, NJ.

John, R.S., von Winterfeldt, D., Edwards, W., 1983. The quality and user acceptance of multiattribute utility analysis performed by computer and analyst. In: Humphreys, P., Svenson, O., Vari, A. (Eds.), Analysing and Aiding Decision Processes. North-Holland, Amsterdam, pp. 301–319.

Kolers, P.A., Palef, S.R., 1976. Knowing not. Memory & Cognition 4, 553–558.

Koriat, A., Goldsmith, M., 1996. Monitoring and control process in the strategic regulation of memory accuracy. Psychological Review 103, 409–517.

LaPorte, R.E., Nath, R., 1976. Role of performance goals in prose learning. Journal of Educational Psychology 68, 260–264.

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

Mazzoni, G., Cornoldi, C., 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., Shimamura, A.P., 1994. Metacognition: Knowing About Knowing. MIT Press, Cambridge, MA.

Miller, G.A., Galanter, E., Pribram, K.H., 1960. Plans and The Structure of Behavior. Holt, New York.

Nelson, T.O., 1992. Metacognition: Core reading. Allyn & Bacon, Needham Heights, MA.

Nelson, T.O., 1993. Judgments of learning and the allocation of study time. Journal of Experimental Psychology: General 122, 269–273.

Nelson, T.O., Dunlosky, J., 1991. When people's judgements of learning (JOLs) are extremely accurate at predicting subsequent recall: The "delayed-JOL effect." Psychological Science 2, 267–270.

Nelson, T.O., Dunlosky, J., 1994. Norms of paired-associate recall during multitrial learning of Swahili-English translation equivalents. Memory 2, 325–335.

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, & Cognition 14, 676–686.

Nelson, T.O., Narens, L., 1990. Metamemory: A theoretical framework and new findings. In: Bower, G.H. (Ed.), The Psychology of Learning and Motivation, vol. 26. Academic Press, New York, pp. 125–141.

Powers, W.T., 1973. Behavior: The Control of Perception. Aldine, Chicago, IL.

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.

Schunk, D.H., Zimmerman, B.J., 1994. Self-regulation of learning and performance. Lawrence Erlbaum, Hillsdale, NJ.

Speer, J.R., Flavell, J.H., 1979. Young children's knowledge of the relative difficulty of recognition and recall memory tasks. Developmental Psychology 15, 214–217.

Thiede, K.W., 1996. The relative importance of anticipated test format and anticipated test difficulty on performance. Quarterly Journal of Experimental Psychology 49A, 901–918.

Wright, G., 1984. Behavioral Decision Theory. Sage, Beverly Hills, CA.