Cognitive Science 30 (2006) 1–16

Metacognitive Control and Optimal Learning

Lisa K. Son[a], Rajiv Sethi[b]

[a]*Department of Psychology, Barnard College*
[b]*Department of Economics, Barnard College*

Received 24 October 2005; received in revised form 24 March 2006; accepted 29 March 2006

Abstract

The notion of optimality is often invoked informally in the literature on metacognitive control. We provide a precise formulation of the optimization problem and show that optimal time allocation strategies depend critically on certain characteristics of the learning environment, such as the extent of time pressure, and the nature of the uptake function. When the learning curve is concave, optimality requires that items at lower levels of initial competence be allocated greater time. On the other hand, with logistic learning curves, optimal allocations vary with time availability in complex and surprising ways. Hence there are conditions under which optimal strategies will be relatively easy to uncover, and others in which suboptimal time allocation might be expected. The model can therefore be used to address the question of whether and when learners should be able to exercise good metacognitive control in practice.

Keywords: Metacognition; Learning; Optimization; Time allocation

1. Introduction

One of the more important and frequent resource allocation problems faced by decision makers in daily life is the allocation of time across competing activities or tasks. An important example of this arises when, in the process of study, learners are confronted with the problem of deciding how to allocate time to a variety of items. This requires that the learner make ongoing judgments regarding the extent to which individual items have been learned and, based on these judgments, to control subsequent allocations of time. These two components—monitoring and control—constitute the general framework of metacognition (Nelson & Narens, 1990, 1994). The processes of monitoring and control result in the allocation of study time, spacing decisions, and testing decisions, in addition to other strategies of study. In this article, we focus on the metacognitive control of time allocation.

Correspondence should be addressed to Lisa K. Son, Department of Psychology, Barnard College, 3009 Broadway, New York, NY 10027. E-mail: lson@barnard.edu

There are a number of plausible ways in which a given amount of available time can be allocated to a set of to-be-learned items. For instance, one might spend most time on those items that are judged to be the most difficult or furthest away from a learned state (Dunlosky & Hertzog, 1998). Alternatively, time may be allocated disproportionately to items of intermediate difficulty, with more challenging items receiving attention only when time pressure is not binding (Metcalfe, 2002). In these and other related theories, the idea of optimality has often been invoked, but without the development of a rigorous model of optimizing behavior. This is the gap in the literature that we seek to fill.

One of the advantages of an explicit model of optimal time allocation is that it makes transparent the manner in which allocations depend on structural characteristics of the learning environment, such as the general shape of the learning curve. Without a precise formulation of the problem faced by the learner, and a detailed characterization of optimal behavior, it is impossible to address the question of the effectiveness of study strategies. For instance, speaking of the individuals in their series of experiments, Metcalfe and Kornell (2005) wrote:

> 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. (p. 476)

Such questions of efficacy can only be addressed once formal optimizing models of the kind developed here have been explored.

Whether or not a particular time allocation strategy is optimal will depend on two key aspects of the learning environment. One is the shape of the learning curve or uptake function, which describes how investments of time result in increases in competence. The other is the set of goals or objectives of the learner. Two different types of learning curves are considered here. In the case of diminishing returns, the slope of the learning curve decreases continuously as higher levels of competence are attained. Hence investments of time raise competence most sharply at low levels of competence. Alternatively, one might have S-shaped or logistic learning curves that have increasing slopes at the lowest levels of initial competence but decreasing slopes once learning has proceeded beyond some level. For certain learner objectives, the qualitative properties of the optimal allocation are highly sensitive to the general shape of the learning curve. Under diminishing returns optimality requires that the items that are initially least well learned receive the greatest time allocations. This is regardless of the extent of time pressure. With logistic learning curves, on the other hand, a smaller and easier set of items are studied under high time pressure. Successively more difficult items receive attention as time pressure eases. A key finding is that with logistic learning curves, optimal allocations vary with time pressure in a manner that is discontinuous and nonmonotonic.

There now exists a considerable body of evidence dealing with time allocation strategies, as well as theories of metacognitive control. We examine later the extent to which observed learning behavior, and the theoretical models that have been developed to explain it, are consistent with optimal time allocation. Moreover, although the focus of this article is on the allocation of time across competing activities, the approach taken here can be applied to other settings. The general question of the efficacy of learning processes in furthering individual goals is of broad interest, and formal models of optimization provide a framework within which such questions of efficacy can be fruitfully addressed.

2. A model of time allocation

When learning, an individual is confronted with a set of items, and has a specific amount of total time to allocate to his or her study. At the outset each item will be at some initial level of learning, which we can identify with a point on the item's learning curve. This is a function that describes the manner in which investments of time result in increased competence. What might these functions look like? Because the level of competence that can be attained in any given task is bounded, learning curves must eventually plateau. This leaves two plausible types of learning curve, as shown in Fig. 1. The curve on the left represents learning under conditions of diminishing returns, where the slope of the curve is steepest at the outset, and flattens as learning progresses. That on the right represents learning that is slow at the outset, increases most rapidly at intermediate levels of time investment, and eventually plateaus. The points *A* and *B* represent initial and final levels of learning, based on some allocation of study time.

Learning curves with diminishing returns are consistent with exponential, hyperbolic, square-root, and power functions, and have been commonly been found in a wide variety of laboratory situations (see, e.g., Anderson & Schooler, 1991, for the case of practice effects). Many have argued, however, that the time scale on which laboratory learning occurs is too short to capture the relevant uptake functions in everyday learning. Newell, Liu, and Mayer-Kress (2001), for instance, noted that the "number of practice trials or the duration of the practice period for the assessment of learning curves has been … quite limited in relation to the realities of the performance of everyday activities" (p. 59). It is conceivable, therefore, that laboratory learning curves in fact capture only the upper portion of an S-shaped learning curve, which more accurately describes learning over a longer time scale (Fischer & Pipp, 1984). Such S-shaped learning curves are predicted by the theoretical models of Hull (1943), Klopf (1988), van Geert (1991), and Newell et al. (2001).

In general, S-shaped uptake functions are plausible for the learning of complex skills. Empirical evidence for such functions have been found in language acquisition (Rice, Wexler, &

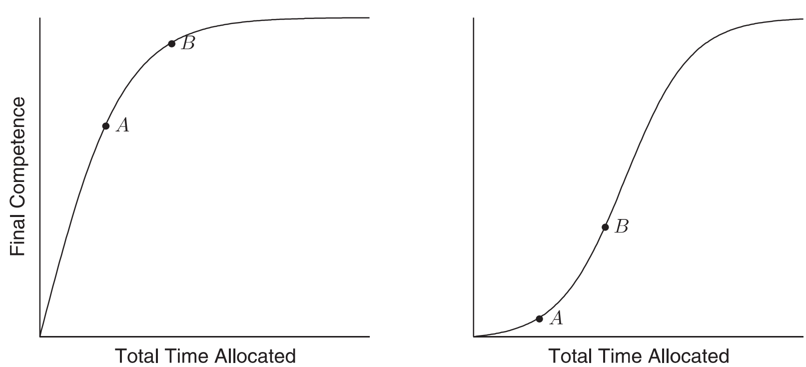

Fig. 1. Examples of learning curves.

Hershberger, 1998), sequence learning (Noble, 1957), motor learning (Newell et al., 2001), and conditioning (Klopf, 1988). For example, when investigating the learning of tenses in both normal and language-impaired children, "inspection of the individual curves shows slow growth at the beginning … followed by rapid acceleration, and then a final period of leveling off" (Rice et al., 1998, p. 1425). Noble (1957) reported results for sequence learning that show curves that are "skewed and S-shaped" (p. 247). Frey and Sears (1978) observed that acquisition curves in conditioning "are typically S-shaped, with a period of positive acceleration followed by one of negative acceleration" (p. 324).

Given a particular set of to-be-learned items, each with its own learning curve, how should the total available time be allocated to best meet the learner's objectives? Let n denote the total number of items to be learned, and T the time available to be allocated. Prior to the allocation of any study time, each of the items will be at some initial level of learning. Let a_i denote this initial level of competence for each item i, where $i = 1, 2, \ldots, n$. This level of competence is the outcome of some prior time allocation s_i to the study of the item. In this case, $a_i = f_i(s_i)$, where f_i is the uptake function or learning curve for item i. Note that we allow for different items to have different learning curves.

Suppose an item is at a point A on its learning curve (as in Fig. 1), where the coordinates of A are (s_i, a_i). Now let t_i denote the amount of additional time allocated to the item. This takes item i to some point B on its learning curve, where the coordinates of B are $(s_i + t_i, b_i)$, and $b_i = f_i(s_i + t_i)$. If no time is left unallocated, we must have

$$T = \sum_{i=1}^{n} t_i .$$

The total available time T identifies a range of feasible time allocations $(t_1, \ldots, t_n)$ among which the learner can choose. Each such choice results in some distribution of final competences $(b_1, \ldots, b_n)$ where b_i is the final competence attained for item i.

The optimal allocation of time across items will depend both on the shapes of the learning curves, and on how various distributions of final competence are valued by the learner. To allow for the possibility that some items are more highly valued than others, we assume that the learner wishes to maximize a weighted average of competences $\sum_{i=1}^{n} w_i b_i$. Here the weight w_i measures the importance to the decision maker of achieving a high level of performance in task i. If all tasks are considered to be equally important, then we have the special case $\sum_{i=1}^{n} b_i$. Here the decision maker simply wants to maximize his or her aggregate score. Note that there are plausible learner objectives that are not explicitly considered here (e.g., learners might care about the minimum competence across all items, the maximum competence of any one item, or the number of items above some specified threshold).

Items may receive greater weight for a variety of reasons including the extent to which the learner finds them intrinsically pleasant or unpleasant. Such motivations may be very important: Judgments of interest have been shown to be positively correlated with investments of time (Son & Metcalfe, 2000). Alternatively, in experimental settings, the weight placed on an item might be manipulated by the experimenter. For instance, Dunlosky and Thiede (1998) achieved this by assigning 10 points for learning some items and 1 point for learning others (see also Le Ny, Denhiere, & Le Taillanter, 1972). Unfortunately instructions given in most experiments are too vague for an observer to make precise inferences about learner objectives.

For any given valuation, optimal behavior will depend on the shape of the uptake function. We consider first the case of diminishing returns, and then examine logistic learning curves.

3. Diminishing returns

Suppose that the learning curves for all items are increasing, so the first derivative $f_i' > 0$ for all i. Suppose further that the curves are concave, so the second derivative $f_i'' < 0$. This is the case of diminishing returns. We now show how optimal allocations vary with total available time in the special case of two items with equal value and identical learning curves. We then consider the more general case of multiple items with varying values and different learning curves.

Suppose that each of two items has the same value ($w_1 = w_2$) and uptake function ($f_1 = f_2$) and that the initial competencies for the items are a_1 and a_2, respectively, with $a_2 > a_1$ (see Fig. 2). Corresponding to these initial competencies are the historical times s_1 and s_2, such that $a_i = f(s_i)$. When learners wish to maximize the sum of competencies, standard results from the theory of classical optimization (see, e.g., Intriligator, 1971) imply that if both items are allocated positive amounts of time, then

$$f'(s_1 + t_1) = f'(s_2 + t_2) \tag{1}$$

In other words, the *marginal returns* to time allocated must be equalized across the two items. This makes sense: If it were *not* the case that Condition 1 were satisfied, then the decision maker could raise the sum of competencies by shifting time away from the item with the lower marginal return and toward the item with the higher marginal return. Because no such profitable reallocations can be possible at an optimal allocation, marginal returns must be equalized across items. Furthermore, because $f'' < 0$ at all points on the learning curve, Condition 1 can only be satisfied if $s_1 + t_1 = s_2 + t_2$, and hence $b_1 = b_2$. That is, if both items receive positive time allocations, then they must be brought to the same point on the common learning curve, as in Fig. 2. Hence the optimizing decision maker should equalize levels of final competence across items, and devote most resources to those items initially at the lowest levels of competence.

What if there is insufficient time available to achieve the equality of slopes required by Condition 1? If total available time T is less than $s_2 - s_1$ the optimal allocation requires that all time be devoted to the initially less well learned item. This follows from the fact that the marginal returns are higher at lower levels of initial competence. Fig. 3 describes how optimal allocations vary for a range of values of T. When total available time is scarce, only Item 1 is studied, but once T is sufficiently large, both items receive attention in such a manner as to maintain equality of slopes and equality of final competence.

These qualitative results generalize to the case of multiple items. When time is scarce, the items that receive positive attention will be those initially at the lowest levels of learning, and any subset of items that receives positive time allocations must be brought to the same point on the learning curve.

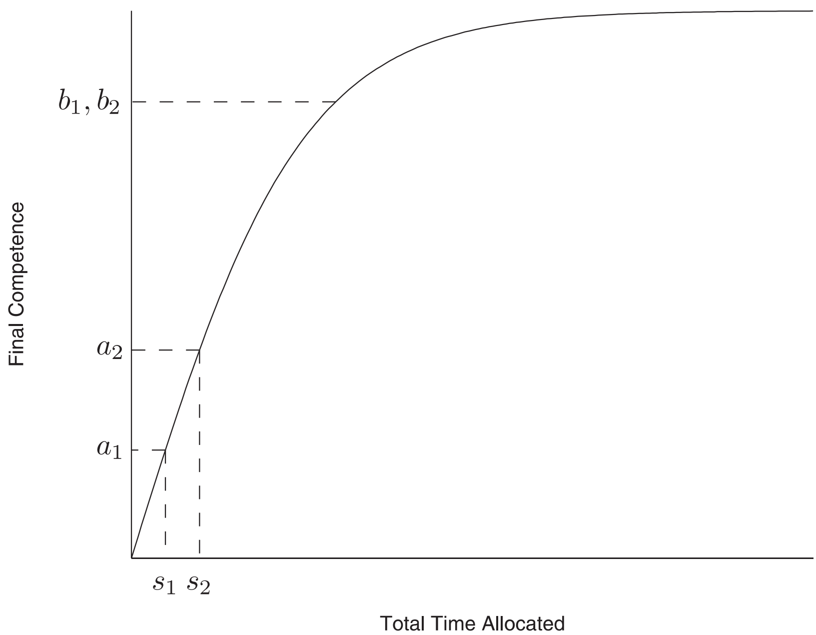

Fig. 2. Diminishing returns with two items.

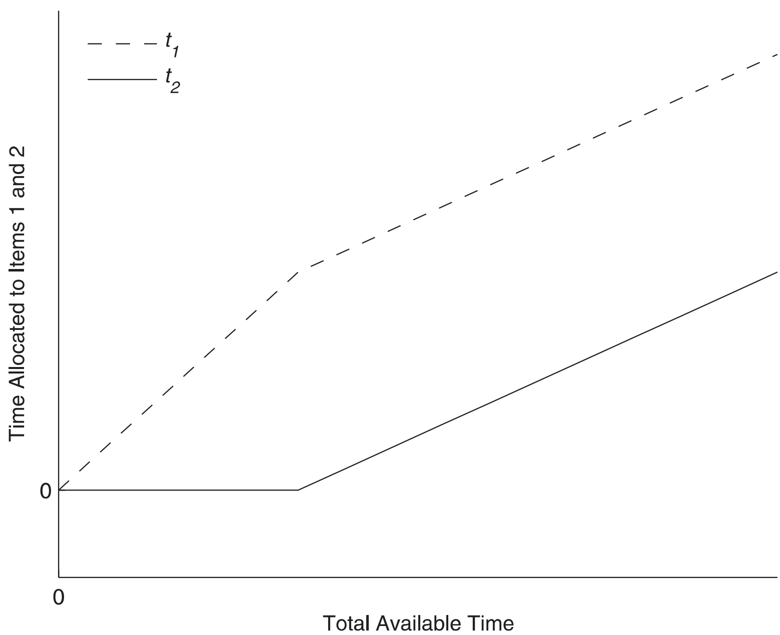

Fig. 3. Optimal allocations under diminishing returns ($a_1 = 0.20$, $a_2 = 0.38$).

The preceding analysis assumes that both the item weights and the learning curves are identical. In practice, items may vary with respect to their intrinsic difficulty as well as the importance placed on them by the learner. We now consider both these possibilities for the special case of exponential learning curves. Suppose that the learning curve for item *i* is given by

$$a_i = f_i(s_i) = 1 - e^{-x_i s_i}, \tag{2}$$

and the learner wishes to maximize the weighted sum of competences $\sum_{i=1}^{n} w_i b_i$.. In this case, items vary both with respect to the difficulty of learning x_i and their value to the learner w_i. Here x_i is a measure of the ease with which the decision maker can raise competence in the particular task *i*. High values of x_i imply that relatively small time investments can have large effects on competence. Similarly, raising competence in some tasks may be more important to the decision maker than doing so in other tasks. This asymmetry would result in unequal weights w_i across tasks.

Now suppose that, given initial competence a_i an additional amount of time t_i is allocated to the task, resulting in a level of final competence b_i. Then

$$b_i = 1 - e^{-x_i(s_i + t_i)} = 1 - e^{-x_i s_i} e^{-x_i t_i} = 1 - (1 - a_i)e^{-x_i t_i} \cdot$$

What is the optimal allocation of time across tasks in this case? As before, optimality requires the equalization of marginal returns to time allocated across the tasks. With different learning curves and item weights, optimality requires that for every pair of items *i* and *j,*

$$w_i f_i'(s_i + t_i) = w_j f_j'(s_j + t_j).$$

Hence the following must hold at any optimal allocation:

$$w_i x_i (1 - a_i)e^{-x_i t_i} = w_j x_j (1 - a_j)e^{-x_j t_j}.$$

This can be rewritten as

$$x_i t_i - x_j t_j = \log\left(\frac{w_i x_i (1 - a_i)}{w_j x_j (1 - a_j)}\right).$$

In the special case of $x_i = x_j$ and $w_i = w_j$ for some pair *i* and *j* (where both tasks have the same learning curve and value to the learner) the above implies that $t_i < t_j$ if and only if $a_i > a_j$. In other words, the optimizing decision maker allocates more time to the task that has a lower level of initial competence. However, if either $x_i \neq x_j$ or $w_i \neq w_j$ then it is possible for the task with higher initial competence to receive more attention. This is especially likely to occur if the intrinsic difficulties of learning the two items are vastly different, or if one item is much more highly valued than the other. Allocating time to the task that requires more time to increase competence can involve laboring in vain (Mazzoni & Cornoldi, 1993; Mazzoni, Cornoldi, & Marchitelli, 1990; Nelson & Leonesio, 1988). Similarly tasks in which raising competence is of greatest value will tend to receive more attention, unless they are at significantly flatter portions of the learning curve.

4. Logistic learning curves

Now suppose that the learning curves are S-shaped, as in the right panel of Fig. 1. Specifically, suppose that there exists some time availability τ such that the slope f' is increasing for time allocations below τ and decreasing above. This implies that $f''(\tau) = 0$, $f''(s_i + t_i) > 0$ for $s_i + t_i < \tau$ and $f'(s_i + t_i) < 0$ for $s_i + t_i > \tau$ as in Fig. 4. In this case, the manner in which optimal allocations vary with total available time is both more complicated and more interesting, even in the simplest case where items have equal values and identical learning curves.

Consider the case of two items and suppose that initial competences a_1 and a_2 are both below $f(\tau)$, as in Fig. 4. That is, we have $s_1 < s_2 < \tau$ so both items are on the segment of the learning curve with increasing slope. It is then possible to show that there exists a threshold value $\hat{T}$ of total available time such that the optimal allocation assigns all available time to Item 2 whenever $T < \hat{T}$. If $T > \hat{T}$ the optimal allocation assigns positive time to both items but with greater time given to Item 1 (see the Appendix for a formal proof of this claim). At the critical value of total time $\hat{T}$, the optimal strategy shifts discontinuously from allocations in which attention is focused exclusively on Item 2 to allocations in which Item 1 receives more attention than Item 2. Crucially, the amount of time allocated to Item 2 decreases in absolute terms as total available time crosses this critical value. Hence Item 2 is brought further away from the plateau of its learning curve at this point. Fig. 5 illustrates this phenomenon. Within each regime, optimal allocations vary smoothly with changes in total time T. At the point of transition between the two regimes, though, there is a jump in the allocations to both items, with t_1 leapfrogging t_2.

The basic findings in the two-item case generalize to the case of multiple items. Optimal allocations in the case of three items are depicted in Fig. 6. As before, Item 1 has the lowest level

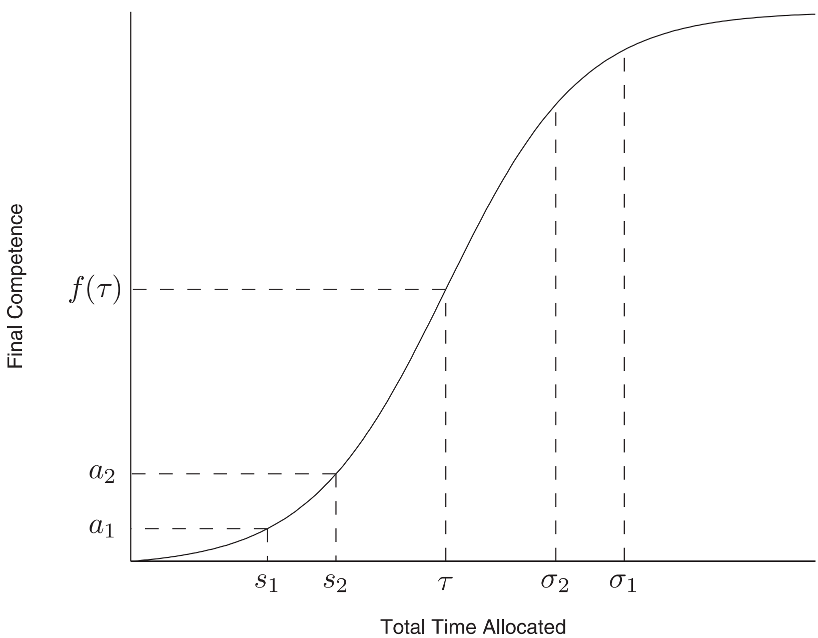

Fig. 4. Two items with initial competence below $f(\tau)$.

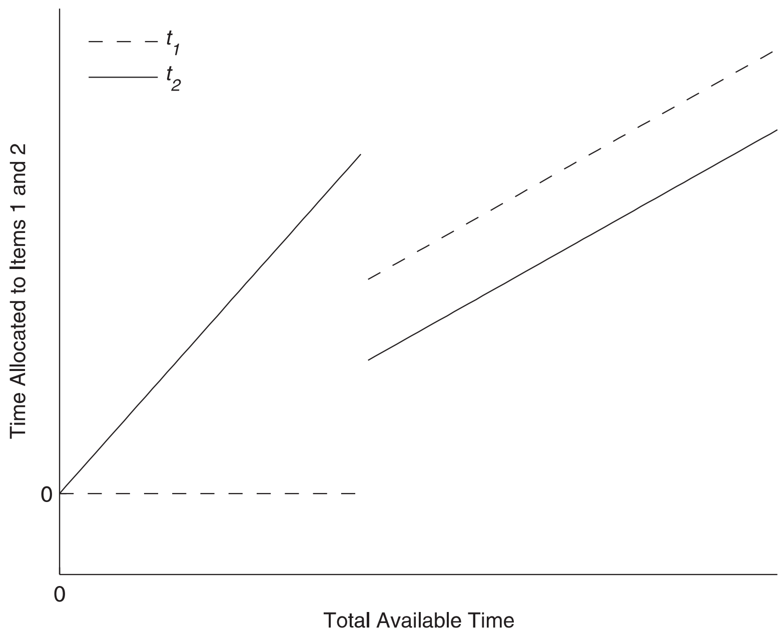

Fig. 5. Optimal allocations with logistic learning ($a_1 = 0.06$, $a_2 = 0.16$).

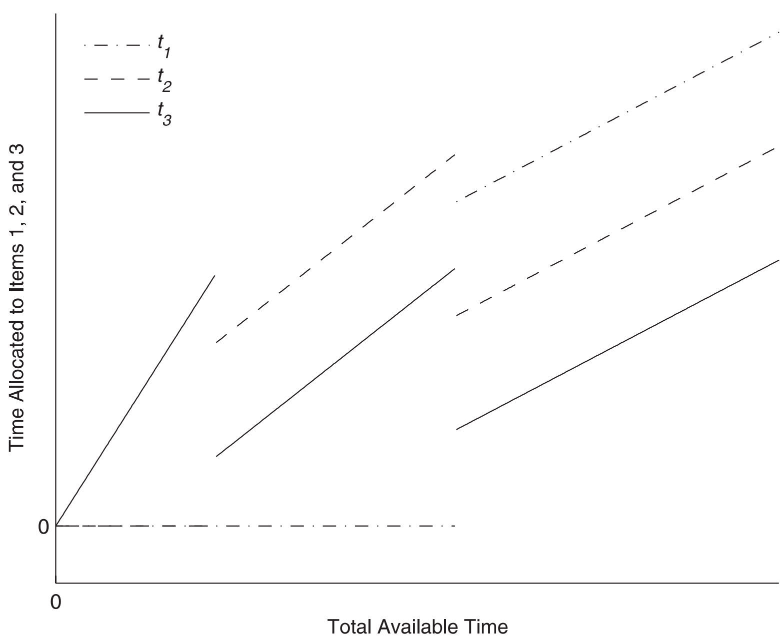

Fig. 6. Optimal allocations with logistic learning ($a_1 = 0.06$, $a_2 = 0.16$, $a_3 = 0.35$).

of initial competence, and Item 3 the highest. In this case two transitions occur, and three regimes may be identified. When time pressure is highest, all attention is devoted to Item 3. As time availability rises, a transition point is reached at which t_3 drops discontinuously and t_2 rises from zero to a point above t_3. Both Items 2 and 3 receive attention throughout this regime, with more time being allocated to the initially less well learned Item 2. As time availability becomes even more abundant, a second transition point is reached at which both t_2 and t_3 drop discontinuously, and t_1 rises from zero to a point above t_2. Within this last regime, all items are given at least some attention, with most attention going to Item 1 followed by Item 2 and then Item 3. As in the case of two items, optimal allocations vary smoothly within regimes, but shift in a discontinuous and nonmonotonic manner at transition points.

The preceding discussion (including Figs. 4–6 and the formal result in the Appendix) are based on the hypothesis that different items are at different levels of initial competence on the same learning curve. Optimality then requires that, for any given level of time availability, the subset of items chosen for study are those closest to a learned state. Within this set of items, however, greatest attention is paid to those items with the lowest levels of initial competence. For reasons discussed in the previous section, these basic conclusions remain intact even when learning curves and value weights differ across items, as long as such differences are not too great.

A limiting case of the logistic learning curve (as the curves become increasingly steep in the intermediate range) is a *step function*. Such functions are entirely flat (at zero competence) until some threshold level of time allocation has been reached, and then jump discontinuously to a fully learned state. In the special case of identical thresholds for each item, the optimal allocation is easily characterized. When total time available is too low to bring any item to its threshold then all allocation strategies are equally ineffective. As time available increases, there comes a point at which it becomes possible for the learner to bring one item (the one with the highest initial competence) to its threshold. It is optimal for the learner to allocate all available time to this item. Further increments in total available time may then be allocated to the item that is now closest to its threshold but still below it. Reasoning in this manner we see that optimal behavior entails allocations to a sequence of tasks, starting with the one with initially highest competence and moving down the lists of tasks in order of initial competence.

Step functions can be quite realistic descriptions of learning in environments where the learner simply needs to reach some threshold level of competence. Consider, for instance, the case of a driving test in which passing requires only that a certain level of errors is not exceeded. If competence is measured as the likelihood of passing, then time allocations within a broad range can reduce errors without raising competence. Uptake functions with diminishing returns simply cannot apply to learning under these conditions.

5. Existing Theory and Evidence

Although a thorough empirical analysis of the model's predictions is well beyond the scope of this article, we here take a first look at the available data from the perspective of the optimization framework. Data on time allocation have typically been collected using the following experimental paradigm. Learners are initially confronted with a list of to-be-learned items (typically words) and asked to assess their judgments of learning for each item. Next, each of the items is presented in sequence, for a duration determined by the learner. Because time constraints are absent, participants are free to allocate as much time as they wish to each item, and the amount of time spent on each item is recorded. Using this method, results have shown that individuals tend to allocate more study time to the judged-difficult items (those further away from a learned state) than to the judged-easy items (Cull & Zechmeister, 1994; Mazzoni & Cornoldi, 1993; Mazzoni et al., 1990; Nelson, Dunlosky, Graf, & Narens, 1994; Nelson & Leonesio, 1988; Thiede & Dunlosky, 1999). Based on the data collected using these methods, the *discrepancy reduction* hypothesis was formulated to describe how people used their metacognitive judgments to control subsequent time allocation (Dunlosky & Hertzog, 1998). According to discrepancy reduction, learners compare the degree of discrepancy between the current state of an item (e.g., unlearned, almost learned, learned, etc.) and their own desired state of learning for that item. They then allocate time disproportionately to those items that are characterized by the highest levels of initial discrepancy.

Our model shows that discrepancy reduction strategies may indeed be optimal, provided that the learning curve is characterized by diminishing returns, and learners maximize the sum of competencies across items. On the other hand, with logistic learning curves, discrepancy reduction strategies are no longer optimal in general. In this case, the set of items chosen for study are those initially closest to a learned state, although within the subset of items selected for study, those furthest from a learned state receive the most attention. Hence, discrepancy reduction holds in the limited sense that it applies only to the subset of items chosen for study. These properties are consistent with the hierarchical model and experimental findings of Thiede and Dunlosky (1999) and Dunlosky and Thiede (2004).

More recently, there have been a number of studies that have challenged the discrepancy reduction view (Metcalfe, 2002; Metcalfe & Kornell, 2003, 2005; Son & Metcalfe, 2000). Son and Metcalfe (2000), for example, examined the effects of time pressure on people's time allocation decisions. Their results showed that people did allocate more time to items that were judged as difficult, but only under conditions in which participants were not time pressured. Under high time pressure conditions, people's time allocation strategies shifted toward allocating time to the judged-easy items. Similarly, Metcalfe (2002) considered three distinct time pressure conditions: low, medium, and high. Under these circumstances, people behaved in a manner consistent with discrepancy reduction only in the low-pressure condition. In the medium-pressure condition, the most time was allocated to items of medium difficulty, and in the high-pressure condition, most time was allocated to items of low difficulty. This led to the development of an alternative approach to time allocation, the *region of proximal learning* model.

The region of proximal learning model is based on the notion that people have a "zone" of learning where they allocate the most study time to items that are not too easy, but not so difficult that the returns on the investments of time are extremely low (Metcalfe, 2002; Metcalfe & Kornell, 2003, 2005). Specifically:

> If there is a range of to-be-learned items, then before anything is learned, the easiest items will be in the proximal learning state and will gain the most from study. Once those easy items are mastered, though, little additional gain would be expected for additional study effort on them, and the region of

> proximal learning should shift to an item set that is more difficult. Learning, then, is reflected in a shift toward study of items of progressively greater difficulty. (Metcalfe, 2002, p. 350)

These findings are interpreted by the authors as reflecting optimal behavior under diminishing returns, but with substantial differences in learning curves across items (Metcalfe & Kornell, 2005). As noted earlier, if some learning curves are much flatter than others, a discrepancy reduction strategy would not generally be optimal, even under diminishing returns.

An alternative interpretation of these findings (from the perspective of our model) is that learning curves are roughly similar across items, but are logistic in shape, with difficult items being at the flat initial segment of the curve, and the easiest items further up and closer to the plateau or learned state. In this case, items of moderate difficulty, which lie on the steepest segment of the learning curve, would receive the highest time allocations under optimal learning. Items that are either too easy or too difficult would receive less time, as investments in these items would result in very small gains. One way to implement such a strategy would be to begin by investing in items of intermediate difficulty until these are brought to flatter portions of the learning curve, and then moving in sequence to increasingly difficult items. This is precisely the pattern reported by Metcalfe and Kornell (2005).

6. Discussion

Existing data in the time allocation literature have commonly been interpreted as indicative of optimal learning (e.g., Metcalfe & Kornell, 2005; Son & Metcalfe, 2000; Thiede & Dunlosky, 1999). The model developed here makes transparent the fact that the properties of optimal allocations are highly sensitive to structural characteristics of the learning environment. Hence strategies that are highly effective with one class of learning curves may be quite ineffective with another. This possibility seems to have been largely neglected in theoretical models of metacognitive control.

The model of optimal time allocation developed here implies a systematic relation between the extent of total available time and the pattern of its use. For instance, when items are comparable with respect to value and rate of learning, those items furthest away from a learned state receive the most attention, regardless of the extent of time pressure under diminishing returns. With learning curves of a logistic type, however, small differences in total available time can result in large qualitative differences in the nature of the optimal strategy. For any given level of time availability, the set of items chosen for study are those closest to a learned state, but within this set, the ones furthest away from a learned state receive the most attention. The non-monotonicity at transition points and the leapfrogging (with respect to time allocation) of initially better learned items by items that are initially less well learned are novel predictions of this model. Hence optimal time allocation strategies depend critically on certain characteristics of the learning environment, such as the extent of time pressure, and the nature of the uptake function.

Under what circumstances might optimal allocations be uncovered by the learner? In practice, critical information about the basic structure of the learning environment (e.g., the general form of the uptake functions) may be unknown. Because cognitive capacity is limited, learners may be forced to adopt simple heuristics or rules of thumb based on limited information, such as the slopes of the learning curves in the immediate vicinity of the current allocation. For instance, consider the simple rule that prompts learners to allocate resources to whichever task has the steepest (value-adjusted) uptake at the current time. This will result in optimal time allocation under diminishing returns, but not under logistic uptake functions. Hence there will generally arise situations in which the predictions of optimizing models fail to accurately match real-world behavior. The idea that optimization models implicitly assume that no resources (cognitive or material) are required to solve optimization problems was emphasized by Simon (1978) and Conlisk (1988). These limitations are particularly evident when considering the time allocation problem, because the formulation and solution of optimization problems is clearly time intensive, and any attempt at optimization runs the risk of squandering the scarce resource. Hence it is not surprising that learners have been found to use simple but ineffective strategies in certain contexts (as in Atkinson, 1972) and adopting effective strategies in others (as in Metcalfe & Kornell, 2003). The optimization model can be used to identify conditions under which behavior should be consistent with optimal choice, and when it is likely to be inconsistent. It can therefore be used to address the question of whether and when learners should be able to exercise good metacognitive control in practice.

One of the main objectives of this article is to explore in a rigorous manner the implications of optimizing behavior under various learning conditions. Because the notion of optimality is often invoked informally in the literature on learning, we consider it useful to provide a precise formulation of the optimization problem and its solution. Nevertheless, several interrelated psychological processes have not been captured in the current model. We have ignored the possibility that competence may decline over time in the absence of reinforcement. Additionally, we have assumed implicitly that the level of final competence attained depends only on the cumulative total time allocated to that item, and not on the manner in which this allocation is spaced over time, or sequenced in relation to allocations to other items. Issues of sequencing and spacing have been shown to be empirically important (Son, 2004), and extensions of the current model in this direction are a priority for future research.

Acknowledgments

We thank John Dunlosky, Stephen Lewandowsky, Janet Metcalfe, and especially Robert Goldstone and Keith Raynor, for helpful comments on earlier versions, and the Institute for Advanced Study at Princeton and the Santa Fe Institute for their hospitality.

References

Anderson, J. R., & Schooler, L. J. (1991). Reflections of the environment in memory. *Psychological Science, 2,* 396–408.

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

Conlisk, J. (1988). Optimization cost. *Journal of Economic Behavior and Organization, 9,* 787–821.

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.

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. 249–276). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

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

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

Fischer, K. W., & Pipp, S. L. (1984). Processes of cognitive development: Optimal level and skill acquisition. In R. J. Steinberg (Ed.), *Mechanisms of cognitive development* (pp. 45–80). San Francisco: Freeman.

Frey, P. W., & Sears, R. J. (1978). Models of conditioning incorporating the Rescorla–Wagner associative axiom, a dynamic attention process, and a catastrophe rule. *Psychological Review, 85,* 321–340.

Hull, C. L. (1943). *Principles of behavior, an introduction to behavior theory.* New York: Appleton-Century.

Intriligator, M. D. (1971). *Mathematical optimization and economic theory.* Englewood Cliffs, NJ: Prentice-Hall.

Klopf, A. H. (1988). A neuronal model of classical conditioning. *Psychobiology, 16,* 85–125.

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

Mazzoni, G., & Cornoldi, C. (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.

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

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. (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.

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.

Newell, K. M., Liu, Y. T., & Mayer-Kress, G. (2001). Time scales in motor learning and development. *Psychological Review, 108,* 57–82.

Noble, C. E. (1957). The length–difficulty relationship in compound trial-and-error learning. *Journal of Experimental Psychology, 55,* 246–252.

Rice, M. L., Wexler, K., & Hershberger, S. (1998). Tense over time: The longitudinal course of tense acquisition in children with specific language impairment. *Journal of Speech, Language, and Hearing Research, 41,* 1412–1431.

Simon, H. A. (1978). Rationality as process and as product of thought. *The American Economic Review Papers and Proceedings, 68,* 1–16.

Son, L. K. (2004). Metacognitively-controlled spacing of study. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 30,* 601–604.

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

Thiede, K. W., & Dunlosky, J. (1999). Toward a general model of self-regulated study: An analysis of selection of items for study and self-paced study time. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 25,* 1024–1037.

Van Geert, P. (1991). A dynamic systems model of cognitive and language growth. *Psychological Review, 98,* 3–53.

Appendix

Proposition

Suppose that the learning curve $f(s+t)$ has the following properties $f' > 0, f''(\tau) = 0, f''(s_i + t_i) > 0$ for $s_i + t_i < \tau$, and $f''(s_i + t_i) < 0$ for $s_i + t_i > \tau$. Suppose also that initial competencies a_1 and a_2 are such that $a_1 < a_2 < f(\tau)$. Then there exists a threshold value $\hat{T}$ of total available time such that the optimal allocation $(t_1, t_2) = (0, T)$ for $T < \hat{T}$. For $T > \hat{T}$, the optimal allocation (t_1, t_2) satisfies $t_1 > t_2 > 0$.

Proof

The optimization problem may be stated as follows: The learner must choose t_1 and t_2 to maximize the sum $b_1 + b_2$, subject to the constraints $t_1 \geq 0$, $t_2 \geq 0$, $t_1 + t_2 = T$. Applying the method of Lagrange multipliers (Intriligator, 1971, chap. 3) to this problem, we get the following Lagrangian

$$\begin{aligned} L &= b_1 + b_2 + \lambda_1 t_1 + \lambda_2 t_2 + \lambda_3 (T - t_1 + t_2) \\ &= f(s_1 + t_1) + f(s_2 + t_2) + \lambda_1 t_1 + \lambda_2 t_2 + \lambda_3 (T - t_1 + t_2) \end{aligned}$$

where λ_1, λ_2, and λ_3 are the Lagrange multipliers. Any maximum must satisfy the first order conditions for optimality:

$$\frac{\partial L}{\partial t_1} = f'(s_1 + t_1) + \lambda_1 - \lambda_3 = 0,$$

$$\frac{\partial L}{\partial t_2} = f'(s_2 + t_2) + \lambda_2 - \lambda_3 = 0.$$

Any maximum in which both t_1 and t_2 are positive satisfies $\lambda_1 = \lambda_2 = 0$ and hence $f'(s_1 + t_1) = f'(s_2 + t_2)$, which is Equation 1 in the text.

First we show that if T is sufficiently small, the optimal allocation is $(t_1, t_2) = (0, T)$. Suppose $T < \min\{s_2 - s_1, \tau - s_2\}$ (see Fig. 4). Then at any feasible (t_1, t_2) we have $f'(s_1 + t_1) < f'(s_2 + t_2)$, so any allocation with $t_1 > 0$ cannot be optimal.

Next we show that if T is sufficiently large, then t_1 and t_2 are both positive. Define $\sigma_1 > s_1$ as the unique point at which the learning curve has the same slope as it does at s_1 (see Fig. 4). That is, $f'(\sigma_1) = f'(s_1)$. Suppose $T \geq \sigma_1$. If $(t_1, t_2) = (0, T)$, then $f'(s_2 + T) < f'(s_1)$, which is inconsistent with optimality. On the other hand, if $(t_1, t_2) = (T, 0)$, then $f'(s_1 + T) \leq f'(s_1) < f'(s_2)$, again inconsistent with optimality. Hence, both t_1 and t_2 must be positive.

Next we show that $(t_1, t_2) = (T, 0)$ is never optimal. Suppose, by way of contradiction, there exists some T such that $(T, 0)$ is optimal. Then $f'(s_1 + T) > f'(s_2)$, and hence $s_2 < s_1 + T < \sigma_2$ (see Fig. 4). But in this case the same value of $b_1 + b_2$ can be attained by setting $(t_1, t_2) = (s_2 - s_1, T - (s_2 - s_1))$. However this cannot be optimal because, in this case, $(s_1 + t_1, s_2 + t_2) = (s_2, s_1 + T)$ and $f'(s_2) < f'(T + s_1)$.

Let $\tilde{T}$ be the largest value of T for which $(t_1, t_2) = (0, T)$ is optimal. Clearly $f'(s_1) \le f'(s_2 + \tilde{T})$. Then for $T > \tilde{T}$, t_1 and t_2 are positive. This implies the equality of slopes condition

$$f'(s_1 + t_1) = f'(s_2 + t_2).$$

If $s_1 + t_1 = s_2 + t_2$ then $t_1 > t_2$. If $s_1 + t_1 \ne s_2 + t_2$, and $t_1 \le t_2$, then $s_1 + t_1 < s_2 + t_2$ and equality of slopes implies

$$f''(s_1 + t_1) > 0 > f''(s_2 + t_2). \tag{3}$$

This condition is inconsistent with a maximum. To see this, note that a necessary second order condition for a maximum is that the following Hessian matrix be negative semidefinite (Intriligator, 1971):

$$H = \begin{pmatrix} \dfrac{\partial^2 L}{\partial t_1^2} & \dfrac{\partial^2 L}{\partial t_1 \partial t_2} \\ \dfrac{\partial^2 L}{\partial t_2 \partial t_1} & \dfrac{\partial^2 L}{\partial t_2^2} \end{pmatrix} = \begin{pmatrix} f''(s_1 + t_1) & 0 \\ 0 & f''(s_2 + t_2) \end{pmatrix}$$

The matrix H can only be negative semidefinite if $f''(s_1 + t_1)$ and $f''(s_2 + t_2)$ are both nonpositive, contradicting Equation 3. Hence $t_1 \le t_2$ is impossible at an optimal allocation in which both t_1 and t_2 are positive.