PSYCHOLOGICAL SCIENCE

Research Article

UTILIZATION OF METACOGNITIVE JUDGMENTS IN THE ALLOCATION OF STUDY DURING MULTITRIAL LEARNING

Thomas O. Nelson,[1] John Dunlosky,[1] Aurora Graf,[1] and Louis Narens[2]
[1]University of Washington and [2]University of California, Irvine

Abstract—*We contrasted several ways that an individual's judgments of learning (JOLs) can be utilized when allocating additional study ("restudy") during the learning of Swahili-English translation equivalents The findings demonstrate how metacognitive monitoring can be utilized to benefit multitrial learning Computer-controlled allocation of restudy based on people's JOLs was equivalent to most people's own allocation of restudy (indicating that the computer algorithm can provide a sufficient account of people's allocation of restudy) and was more effective than a computer-controlled allocation based on normative performance (indicating that people's metacognitive monitoring of idiosyncratic knowledge has functional utility in causal chains for learning)*

Self-monitoring and control are fundamental categories of metacognition and consciousness (Kihlstrom, 1984) Few people nowadays would doubt the importance of self-monitoring as a construct in theories of metacognition and consciousness, and much research has been conducted on factors that affect self-monitoring judgments or the accuracy of those judgments at predicting memory performance (e g , see Nelson, 1992) However, perhaps an even more fundamental issue is whether self-monitoring can have a causal role in the ongoing control of learning This issue is important both because of its implications for psychological theory (e g , in models of self-directed learning) and because of its potential for applications for optimizing learning

Investigations of the effect of metacognitive monitoring on learning either have examined it indirectly via correlational designs (Bizanz, Vesonder, & Voss, 1978, Maki & Berry, 1984) or, if examining it more directly, have failed to find that metacognitive monitoring facilitates learning (Begg, Martin, & Needham, 1992, Mazzoni & Cornoldi, 1993, Mazzoni, Cornoldi, & Marchitelli, 1990, Nelson & Leonesio, 1988) Those negative findings have led some researchers to conclude that metacognitions "are a form of introspective witness, even when they accurately indicate the state of the system, they have no value for memory" (Begg et al , 1992, p 207) This conclusion implies that metacognitive monitoring is an epiphenomenon rather than part of the causal chain for learning One of our major goals was to develop a simple experiment that would demonstrate how people's self-monitoring can be causally efficacious for multitrial learning

A second goal was to test the adequacy of a computerized algorithm for allocating additional study (hereafter "restudy") to various items This algorithm operates only on the input from people's metacognitive monitoring judgments and simulates a specific form of interplay between metacognitive monitoring and control, which according to theory (Nelson & Narens, 1990, especially their Fig 4) might facilitate learning Thus, the present research investigated the entire three-part causal chain of monitoring affecting control affecting learning The main theoretical supposition contained in the algorithm is that more restudy should be allocated to items that are metacognitively judged to be poorly learned than to items judged to be well learned (Nelson & Dunlosky, 1991, Nelson & Narens, 1990) We evaluated the algorithm in two ways (a) as a performance model for facilitating people's learning and (b) as a simulation model that might be sufficient to account for how people utilize the input from their own metacognitive monitoring

Our investigation can be contrasted with earlier research on the optimization of learning (e g , Atkinson, 1972a, 1972b, Groen & Atkinson, 1966) that investigated computerized "response-sensitive strategies" whose input was only the correctness of the subjects' recall responses Concerning the "sufficient history" incorporated in his optimization model, Atkinson (1972b) wrote, "For the model considered in this paper, the sufficient history is [only] the ordered sequence of correct and incorrect responses to a given item plus the number of errors (to other items)" (p 128) The major components of those optimization models were summarized in Figure 1 of Groen and Atkinson (1966), wherein a flowchart of the general paradigm "contains, as special cases, all other programmed instructional techniques currently in vogue" (p 311) However, the paradigm disregarded people's potentially useful discriminations between various items and did not include any metacognitive components

Rather than assuming people to be homogeneous, we explored the potential importance of individuals' idiosyncratic memories Some researchers (Lovelace, 1984, Schneider & Laurion, 1993) have found that individuals can monitor idiosyncratic aspects of their memories and thereby can outpredict group base-rate performance, whereas other researchers (Nelson, Leonesio, Landwehr, & Narens, 1986) have found that people's eventual memory performance can be predicted more accurately by group base-rate information than by their own metacognitive monitoring (elaborated in the Main Findings) Accordingly, our third goal was to explore whether learning differs when the information input to the aforementioned algorithm comes from people's judgments about the idiosyncratic aspects of their memories or from group base-rate information

The task we investigated was people's multitrial learning of Swahili-English translation equivalents (e g , *ardhi–soil*), which are vocabulary items such as those that people learn in foreign-language courses The first phase of the task required all subjects to study and make a judgment of learning (JOL) for every item The second phase was adapted from previous research on metacognitive control (Masur, McIntyre, & Flavell, 1973) and

---

Address correspondence to Thomas O Nelson, Psychology (NI-25), University of Washington, Seattle, WA 98195

consisted of restudy-test trials in which a fixed subset of the items was restudied prior to every test trial on the entire list, testing all items on every test trial allowed us to determine how much of the list had been mastered at each point in the experiment, as described by Atkinson (1972a, p. 927). This task allowed us to determine how effective the various restudy strategies are for attaining mastery. The critical manipulation was the selection of the particular subset of items that would be restudied. The subset was selected by one of the computer algorithms or by the individual subject, as described next.

METHOD

Items, Subjects, and Design

The items were the 36 Swahili-English translation equivalents having the highest proportion of recall in the Nelson and Dunlosky norms (in press, Trial 1). The spread in recall was greater for those items (with the proportion correct ranging from .55 to .15) than for the remaining 64 items in the norms and thereby allowed for substantial discriminability of item difficulty; also, the overall level of recall left ample room to show the effects of learning.

The subjects were 228 undergraduate students from the University of Washington who participated for course credit and were assigned to one of four groups ($n = 57$ per group) by a block-randomization design in which the $i + 1$th subject in a given group was not run until the ith replication was complete. The design contained one between-subjects independent variable with four levels. In the worst-learned-items group, the 18 items designated for restudy were the least well learned according to the subject's JOLs (i.e., the 18 items receiving the lowest JOLs from that person). In the best-learned-items group, the items designated for restudy were the 18 items having the highest JOLs. The normatively-most-difficult-items group was identical to the worst-learned-items group in that the "worst learned" items were selected for restudy, except that the definition of worst learned items was based on subjective reports about idiosyncratic difficulty for the latter group and on objective, normative difficulty (namely, the 18 list items having the lowest probability of recall according to the Nelson & Dunlosky, in press, norms) for the former group. In the self-chosen-items group, each subject chose the particular items that he or she would restudy (elaborated below).

Procedure

First, a study trial on the 36 items occurred at the rate of 4 s/item. Immediately afterward, the subject made a self-paced JOL on every item in response to the cue, "How confident are you that about 10 minutes from now you will be able to recall the second word of the item when prompted with the first word? (0 = definitely won't recall, 20 = 20% sure, 40 , 60 , 80 , and 100 = definitely will recall)." The accuracy of metacognitive monitoring was enhanced by using the stimulus alone (e.g., *ardhi—?*) as the cue for the JOL (Begg et al., 1992; Dunlosky & Nelson, 1992) and by having all groups make delayed JOLs (Dunlosky & Nelson, 1992; Nelson & Dunlosky, 1991) in such a way that the JOLs on the second 18 studied items did not occur until after the JOLs on the first 18 studied items, which in turn did not occur until after the study of all 36 items.

Immediately after the delayed JOL for a given item, the self-chosen-items group made a judgment about whether to allocate restudy to that item. This judgment was cued by the stimulus alone and the query "Would you like to restudy this item?" The subject responded "yes" or "no" as often as he or she liked, and each item requested for restudy was added to the restudy list. After 18 items had been requested for restudy, any additional response of "yes" caused 1 of the items already designated for restudy to be randomly deleted from the restudy list. If a subject requested fewer than 18 items for restudy, the computer randomly selected enough of the unrequested items to reach a total of 18 for the restudy list. Subjects were informed of the 18-item limit in advance, and the computer displayed both the number of items already designated for restudy and the number of items remaining to be judged.

After the judgments, the item order was rerandomized, and a self-paced paired-associate recall test on the 36 items occurred. This test allowed us to assess the equality of the groups at the outset of the experiment. Each stimulus was presented alone, and the subject typed his or her response into the computer (omissions were not allowed). Next, restudy of the 18 restudy items occurred at the rate of 4 s/item, and another paired-associate recall test on all 36 items followed. Then four additional restudy-test cycles occurred (the order of items was rerandomized prior to each cycle), for a total of six test trials. To minimize the role of incorrect spelling, we scored answers as correct whenever the first three letters were correct. No two answers began with the same three letters.

To familiarize the subject with the complete list to help enhance JOL accuracy (Mazzoni et al., 1990) and to have nonfloor recall, two familiarization trials occurred at the outset. During each familiarization trial, a 6-item primacy buffer was presented (not included in the subsequent study or restudy trials), followed by the 36 critical items.

RESULTS AND DISCUSSION

Main Findings

The learning curves for the four groups are shown in Figure 1. As anticipated, recall on the first test trial (prior to any restudy) did not differ across the four groups, $F(3, 224) = 0.15$. By the sixth test trial, however, the four groups differed significantly in their recall, $F(3, 224) = 28.9$, $p < .01$. Tukey post hoc tests were conducted to isolate the differences and yielded the following statistically reliable ordering ($p < .01$ for every inequality): best-learned-items group < normatively-most-difficult-items group < worst-learned-items group ≈ self-chosen-items group.

The finding that recall was worse for the best-learned-items group than for the normatively-most-difficult-items group and for the worst-learned-items group demonstrates that restudy is more effective when allocated to the more difficult items (as identified either by group base-rate information or by people's JOLs, respectively).

The finding that recall was worse for the normatively-most-difficult-items group than for the worst-learned-items group

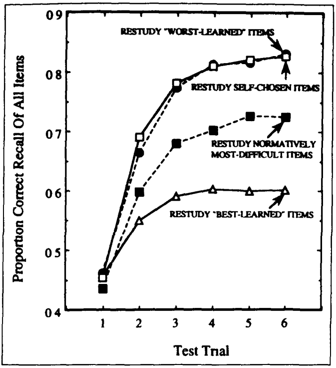

Fig. 1. Mean proportion of all items recalled as a function of trial and group. Restudy occurred on 18 of the 36 items, and all 36 items were tested on every trial.

demonstrates that JOLs can have more functional utility than group base-rate information. This finding about JOLs is different from a related finding about the feeling of knowing (FOK), in which people's subsequent performance on nonrecalled general-information items was predicted more accurately by the normative probability of recall than by the subjects' FOK (Nelson et al., 1986). These outcomes may be reconciled by the fact that FOKs occur only on currently nonretrievable items, whereas JOLs are made on all items, and by the hypothesis that the metacognitive monitoring of retrievable items contributes important idiosyncratic information for the effective allocation of restudy. Additional support comes from the finding (Schneider & Laurion, 1993) that when all items received retrospective confidence judgments, the individuals' recall performance was more highly correlated with those judgments than with normative item difficulty.

The finding that recall was worse for the normatively-most-difficult-items group than for the self-chosen-items group indicates that the allocation of restudy is more effective when designated by individuals than by an algorithm containing the same allocation rule as in the worst-learned-items algorithm but with group base-rate information as the input. This demonstrates that people can use their metacognitions to allocate their restudy effectively. Previous findings, in which people were ineffective at allocating their restudy, may have been due to any of several factors. First, in experiments in which people controlled the duration of self-paced restudy time (Mazzoni & Cornoldi, 1993; Nelson & Leonesio, 1988; Zacks, 1969), the finding of little or no increase in recall for items receiving greater restudy may have been due to the trade-off between extra restudy and extra difficulty of the items (for elaboration, see Nelson, 1993) and to the extra restudy serving functionally as a massed repetition. (Learning is typically no better after massed repetitions than after single presentations of items, e.g., Greeno, 1964.) Second, experiments in which the JOLs were followed by an experimenter-paced restudy of every item (Begg et al., 1992) may not have allowed the subjects ample opportunity to utilize information from their JOLs. To obtain substantial recall advantages from metacognitive activity, people may need the opportunity to choose the items for restudy and may need to have the restudy of a given item be distributed rather than massed (Modigliani & Hedges, 1987).

The finding that recall was equivalent ($p > .90$) for the worst-learned-items group and the self-chosen-items group indicates that the algorithm in the former group is sufficient to account for much of the overall performance in the latter group. Most (but not all—see below) of the people in the self-chosen-items group capitalized on their idiosyncratic information and allocated restudy by a strategy that is functionally similar to the worst-learned-items algorithm.

Fine-Grained Analyses

A priori equivalence of the groups

The accuracy of metacognitive monitoring was assessed by the Goodman-Kruskal gamma correlation (for rationale, see Nelson, 1984). The mean gamma between JOLs and recall on Trial 1 ranged from .88 to .92 and did not differ across groups, $F(3, 222) = 1.67, p > .10$. The mean gamma between the normative probability of recall (from Nelson & Dunlosky, in press) and recall on Trial 1 ranged from .20 to .26 and did not differ across groups, $F(3, 222) = 0.91, p > .10$. For every group, the individual's own JOL accuracy was greater than the accuracy derived from normative probabilities (all ts > 18, all ps < .01), in accord with the finding of greater learning in the worst-learned-items group than in the normatively-most-difficult-items group. Also, correlations between an individual's JOLs and the normative probability of recall (from the Nelson and Dunlosky norms) ranged from .15 to .19 and did not differ across groups, $F(3, 223) = 0.60, p > .10$; these low correlations indicate that substantial idiosyncratic information was being monitored.

Final level of mastery decomposed according to earlier history of recall and restudy

To determine the locus (or loci) of the independent variable's effect on the final level of mastery shown in Figure 1—the probability of correct recall on Trial 6, written as $P(C_6)$—we decomposed $P(C_6)$ into four weighted conditional probabilities.[1] The

---

[1] The conditionalizing terms in these conditional probabilities should be regarded not as causal factors but rather as a way of partitioning the items, thereby allowing us to isolate the subsets of items that were and were not affected by the independent variable. Across groups, the independent variable produced different proportions of items in each subset, and these differences are discussed after consideration of the independent variable's effect on the conditional probabilities.

decomposition is shown by the following equation

$$P(C_6) = p_1\,P(C_6 \mid R \cap C_1) + p_2\,P(C_6 \mid R \cap W_1) + p_3\,P(C_6 \mid N \cap C_1) + p_4\,P(C_6 \mid N \cap W_1), \qquad (1)$$

where $P(C_6 \mid R \cap C_1)$ is the conditional probability of correct recall on Trial 6 for items that were restudied and that were correct on Trial 1, and $P(C_6 \mid N \cap W_1)$ is the conditional probability of correct recall on Trial 6 for items that were not restudied and that were wrong on Trial 1 The weighting factor p_i is the proportion of all items in the list that were in the denominator of the ith conditional probability in Equation 1 (e g , p_1 is the proportion of all items that were restudied and were correct on Trial 1) Notice that $p_1 + p_2 + p_3 + p_4 = 1$, and because exactly 50% of the items were restudied in this experiment, $p_1 + p_2 = 5$ and $p_3 + p_4 = 5$

The independent variable's effect on each conditional probability in the decomposition is shown in a grouped dot chart (Cleveland, 1985, p 151) in Figure 2 Statistical analyses showed no group differences on $P(C_6 \mid R \cap C_1)$ or on $P(C_6 \mid R \cap W_1)$, $F(3, 208) = 0\ 71$ and $F(3, 215) = 1\ 19$, respectively Therefore, the independent variable's substantial effect on the final level of mastery was not due to group differences in the likelihood of being correct on restudied items Similarly, the independent variable had only a small effect on the likelihood of being correct on nonrestudied items On $P(C_6 \mid N \cap C_1)$, the only significant difference ($F[3, 209] = 10\ 03$) was that the best-learned-items group did worse than the other three groups (all $ps < 01$), which did not differ from each other (all $ps > 6$), on $P(C_6 \mid N \cap W_1)$, the only significant difference ($F[3, 200] = 3\ 69$) was that the best-learned-items group did worse than the worst-learned-items group ($p < 01$) Therefore, the group differences in the conditional probabilities of being correct are not sufficient to account for the overall differences in the final level of mastery defined by the ordering reported in the first paragraph of the Main Findings section

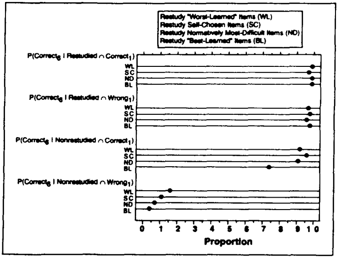

Fig. 2 Grouped dot chart (Cleveland, 1985) of each group's mean proportion of items in each conditional probability of the decomposition (see Equation 1 in the text) of the final level of mastery shown in Figure 1

Instead, the largest effect on the final level of mastery arose from the relative proportions of the different kinds of items that were restudied The four groups differed significantly on each of the weighting factors (p_is) reported at the bottom of Table 1 ($F[3, 224] > 42\ 00$ for each p_i) Because the maximum possible value of each p_i is 5 (i e , exactly 18 of the 36 items were restudied), the observed values across groups span a large portion (almost half) of the possible range for each p_i For both p_2 and p_3, the statistically significant differences across the groups' means in Table 1 mimicked those in the final level of mastery (see Fig 1) Although there was no significant difference between the worst-learned-items group and the self-chosen-items group ($p > 8$ for both p_2 and p_3), all other pairwise differences were significant (all $ps < 01$ for both p_2 and p_3) [2]

In summary, although there were small effects of the independent variable on the intertrial retention of nonrestudied items (third conditional probability in Fig 2) and on the spontaneous recovery of nonrestudied items (fourth conditional probability in Fig 2), the major locus of the effect of the independent variable on the final level of mastery is the groups' differences in the proportion of restudy allocated to items that were initially incorrect (i e , p_2) in combination with the very different levels of eventual recall for initially incorrect items that subsequently were restudied versus nonrestudied (i e , second vs fourth conditional probability, respectively, in Fig 2)—a configural effect The way in which this configural effect occurred can be seen by referring to Equation 1 A large p_1 and small p_3 yielded only a minor advantage over a small p_1 and large p_3, because $P(C_6 \mid R \cap C_1)$ was similar in value to $P(C_6 \mid N \cap C_1)$ (see Fig 2), however, a large p_2 and small p_4 yielded a major advantage over a small p_2 and large p_4, because $P(C_6 \mid R \cap W_1)$ was very much greater than $P(C_6 \mid N \cap W_1)$, which is the largest difference in Figure 2

The subsets of items receiving different JOLs

The second column of Table 1 shows the mean proportion of items that the subjects assigned to each category of JOL The tendency was to assign JOLs more toward the extremes of the JOL scale than equally across all categories (which is typical of delayed JOLs, in contrast to immediate JOLs, Dunlosky & Nelson, in press), but nonetheless the assignment was more finely graded than all-or-none

The next six columns in Table 1 show the mean proportion of correct recall on each trial for items in each category of JOL Items assigned by subjects to the category of JOL = 100% had a high likelihood of being correct on every trial, regardless of whether they were restudied and regardless of group By contrast, the items originally assigned to the lowest categories of JOLs began with low probabilities of recall on Trial 1, and the probabilities of recall increased dramatically in the groups that

---

[2] A perfect inverse relationship existed between the values of p_1 and p_2 and the values of p_3 and p_4 Accordingly, a given group's allocation of restudy to items that had been correct versus wrong on Trial 1 can be seen either in a comparison of the group's values of p_1 and p_2 in Table 1 or in terms of easily computed conditional probabilities, where $P(R \mid C_1) = p_1/(p_1 + p_3)$ and $P(R \mid W_1) = p_2/(p_2 + p_4)$

Table 1 *Proportion of items receiving each judgment-of-learning (JOL) rating, recall performance, and distribution of restudy for each group*

Category of JOL (%)	Mean proportion of items receiving each JOL	Mean proportion of correct recall on each trial						Mean proportion of items in each JOL category that were restudied[a]
		1	2	3	4	5	6	
		Worst-learned-items group						
0	32	09	54	78	85	90	91	94
20	17	22	56	67	72	71	72	62
40	08	37	63	67	65	64	67	43
60	06	60	66	62	67	67	68	28
80	11	78	77	79	80	78	82	25
100	26	90	86	88	88	86	86	04
		Best-learned-items group						
0	33	09	13	15	16	15	15	07
20	17	18	44	50	52	53	53	41
40	06	48	74	79	84	82	84	66
60	05	64	84	87	91	88	91	66
80	10	87	95	97	97	97	95	79
100	30	91	97	98	98	99	98	95
		Normatively-most-difficult-items group						
0	33	06	32	47	52	54	53	52
20	19	20	50	62	64	66	66	58
40	07	30	49	68	67	69	68	50
60	06	52	69	76	79	82	80	57
80	08	70	79	86	86	86	85	49
100	28	91	93	94	92	95	93	43
		Self-chosen-items group						
0	26	05	49	64	69	70	70	69
20	20	18	64	75	75	78	76	72
40	09	29	72	82	87	86	88	73
60	08	51	68	74	76	77	76	46
80	12	75	85	85	84	84	85	34
100	26	92	91	92	92	93	93	11

[a] Collapsed across JOL categories, the distribution of restudy in terms of the weighting factors in Equation 1 was as follows $p_1 = 11$, $p_2 = 39$, $p_3 = 36$, and $p_4 = 14$ for the worst-learned-items group, $p_1 = 35$, $p_2 = 15$, $p_3 = 11$, and $p_4 = 39$ for the best-learned-items group, $p_1 = 19$, $p_2 = 31$, $p_3 = 25$, and $p_4 = 25$ for the normatively-most-difficult-items group, and $p_1 = 12$, $p_2 = 38$, $p_3 = 34$, and $p_4 = 16$ for the self-chosen-items group

restudied those items These patterns, taken together with those in Figure 2, confirm that an important goal of self-directed learning is to allocate restudy so as to make nonrecalled items recallable while maintaining a low likelihood that correctly recalled items will subsequently become nonrecalled

The final column of Table 1 shows the mean proportion of items that received restudy in each category of JOL The differences among the three groups whose restudy was allocated by the computer are due to the algorithms described earlier

The self-chosen-items group

For the self-chosen-items group, the restudy pattern in the final column of Table 1 is most similar to that for the worst-learned-items group However, because of the random deletion of items from the self-chosen-items subjects' restudy lists whenever more than 18 items were requested for restudy, that group's pattern in Table 1 does not show a monotonic decrease across the first three JOL categories This deletion occurred frequently (e g , the median proportion of items requested for restudy was 100% of the items in the first three JOL categories and 75% of the items in the fourth JOL category) Most of the gamma correlations between JOL and requested restudy were close to the corresponding gamma of −1 0 from the worst-learned-items algorithm The median gamma was − 99[3] This

---

[3] The mean gamma was − 83, suggesting that not every subject in the self-chosen-items group followed the worst-learned-items algorithm Two subjects had gammas of 90 and 91, and they reported on a postexperimental questionnaire that they had purposely selected for restudy the items they remembered best, indicating a strategy function-

result confirms that most (but not all) of the subjects in the self-chosen-items group were utilizing their JOLs in a way that is functionally identical to the worst-learned-items algorithm

CONCLUSIONS

Although previous research demonstrated that people's utilization of their monitoring of memory can be ineffective for facilitating subsequent learning, the present findings provide a clear and simple demonstration that sometimes such utilization is effective There is now a need to develop richer theories of metacognition to delineate the domains in which the metacognitive utilization of monitored information is or is not effective for facilitating self-directed learning

As a simulation model, the algorithm for allocating restudy in the worst-learned-items group provides a plausible account of how most (but not all) people utilize input from their own metacognitive monitoring Restudy is allocated more to items that people judge to be poorly learned than to items they judge to be well learned As a performance model, the algorithm for the worst-learned-items group facilitated learning more than did either of the other computer-controlled algorithms

The restudy algorithms for the worst-learned-items group and the normatively-most-difficult-items group differed only in their input (namely, idiosyncratic information from JOLs vs group base-rate information from norms, respectively) Those algorithms are similar to what Karush and Dear (1966, Theorem 1) proved is an optimal strategy, except that Karush and Dear's unit of analysis was a single moment, whereas ours is a block of restudy trials The procedure for the worst-learned-items group differs in three additional ways from related procedures investigated earlier (Atkinson, 1972b, Groen & Atkinson, 1966) First, the input to the algorithm in those earlier investigations came from all-or-none models of learning that ignored the role of short-term memory (Atkinson, 1972b, p 128), whereas the input we examined was either from the person's delayed JOLs (wherein short-term memory factors are taken into account, see Dunlosky & Nelson, in press, and Nelson & Dunlosky, 1991) or from the normative probability of correct recall Second, the earlier investigations used the anticipation method of paired-associate learning, whereas we used the study-test method The latter method may have some advantage in eliminating massed presentations and the attendant complications produced by short-term memory (see Groen & Atkinson, 1966, p 319) Third, the input to the present algorithms was a single-stage decision process (Groen & Atkinson, 1966) that occurred only at the end of the first study trial A natural next step would be to explore multistage decision processes (analogous to the ones used in Atkinson, 1972b), in which the decision about the allocation of restudy would be made anew after each trial and would be based on JOLs made after each of the study or restudy trials

Atkinson (1972a) investigated learner-controlled instruction but not the possibility of using the learner's JOLs as input to a computerized algorithm One of his last remarks about the topic was, "There obviously is a place for the learner's judgments in making instructional decisions However, using the learner's judgment as one of several items of information in making an instructional decision is quite different from proposing that the learner should have complete control" (p 930) Perhaps people's JOLs can be included as input to optimization models of learning so as to replace or augment the input from traditional models of learning Additional research is needed to determine which aspects of restudy should be allocated by a computer versus the learner, when the goal is to optimize self-directed learning The present findings may also be useful when a computer is unavailable, so that the individual is forced to allocate his or her own restudy during learning

Acknowledgments—This research was supported by Grant R01-MH32205 and a career development award (K05-MH1075) from the National Institute of Mental Health to the first author We thank Harry Bahrick, Jim Greeno, and Bill Talbott for their comments on an earlier version of the article

REFERENCES

Atkinson, R C (1972a) Ingredients for a theory of instruction *American Psychologist, 72* 921–931

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

Begg, I M , Martin L A , & Needham, D R (1992) Memory monitoring How useful is self-knowledge about memory? *European Journal of Cognitive Psychology 4* 195–218

Bizanz, G L , Vesonder, G T & Voss, J F (1978) Knowledge of one's own responding and the relation of such knowledge to learning *Journal of Experimental Child Psychology 25*, 116–128

Cleveland, W S (1985) *The elements of graphing data* Pacific Grove, CA Wadsworth

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

Dunlosky, J , & Nelson T O (in press) Does the sensitivity of judgments of learning (JOLs) to the effects of various study activities depend on when the JOLs occur? *Journal of Memory & Language*

Greeno J G (1964) Paired-associate learning with massed and distributed learning of items *Journal of Experimental Psychology 67* 286–295

Groen, G , & Atkinson, R C (1966) Models for optimizing the learning process *Psychological Bulletin 66* 309–320

Karush, W , & Dear, R E (1966) Optimal stimulus presentation strategy for a stimulus sampling model of learning *Journal of Mathematical Psychology 3* 19–47

Kihlstrom, J (1984) Conscious subconscious, unconscious A cognitive perspective In K S Bowers & D Meichenbaum (Eds) *The unconscious reconsidered* New York Wiley

Lovelace, E A (1984) Metamemory Monitoring future recallability during study *Journal of Experimental Psychology Learning Memory and Cognition 10* 756–766

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

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

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

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

Modigliani, V , & Hedges, D G (1987) Distributed rehearsals and the primacy

---

ally like that of the best-learned-items algorithm Five subjects had intermediate gammas, unlike either the worst-learned-items algorithm or the best-learned-items algorithm, and 1 subject had an indeterminate gamma because he requested restudy of every item However, the overwhelming majority—49 of the 57 self-chosen-items subjects—had gammas between − 86 and − 1 inclusive

effect in single-trial free recall *Journal of Experimental Psychology Learning, Memory, and Cognition 13* 426–436

Nelson, T O (1984) A comparison of current measures of the accuracy of feeling-of-knowing predictions *Psychological Bulletin 95*, 109–133

Nelson, T O (1992) *Metacognition Core readings* Boston Allyn & Bacon

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 judgments of learning (JOLs) are extremely accurate for predicting subsequent recall The "delayed JOL effect " *Psychological Science*, 2, 267–270

Nelson, T O , & Dunlosky J (in press) Norms of paired-associate recall for recently studied Swahili-English translation equivalents *Memory*

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 , Leonesio, R.J , Landwehr, R S , & Narens, L (1986) A comparison of three predictors of an individual's memory performance The individual's feeling of knowing versus the normative feeling of knowing versus base-rate item difficulty *Journal of Experimental Psychology Learning, Memory, and Cognition, 12* 279–287

Nelson, T O , & Narens, L (1990) Metamemory A theoretical framework and some new findings In G H Bower (Ed), *The psychology of learning and motivation* (Vol 26, pp 125–173) San Diego Academic Press

Schneider, S L , & Laurion, S K (1993) Do we know what we've learned from listening to the news? *Memory and Cognition, 21*, 198–209

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

(RECEIVED 7/29/93, REVISION ACCEPTED 2/14/94)

Forthcoming Papers

Feature Reviews

Telling What We Know
Robert J Krutcher

Thinking Aloud Insights into Information Processing
John W Payne

Commentary to Feature Review

The Proper Protocol Validity and Completeness of Verbal Reports
Timothy D Wilson

Research Article

Must Figure-Ground Organization Precede Object Recognition? An Assumption in Peril
Mary Peterson and Bradley Gibson

Research Report

Learning and Intermodal Transfer of Information in Newborns
Kelley L Kaye and T G R Bower

Technical Commentary

Localization of a Neural System for Error Detection and Compensation
Stanislas Dehaene, Michael I Posner, and Don M Tucker