Large effects on long-term retention after standard list learning vs. adjusted learning

THOMAS O. NELSON, ARTHUR P. SHIMAMURA, and R. JACOB LEONESIO
University of Washington, Seattle, Washington 98195

Prior research using the method of adjusted learning has focused on the method's potential relative to the standard list learning method for minimizing retention differences among items. Concomitantly, there is less overlearning of individual items with the adjusted learning method. Although prior research has reported a retention difference favoring the standard list learning method, our research demonstrates that the magnitude of this difference can be surprisingly large. The results presented here from 11 experiments show retention to be more than 100% greater after standard list learning than after adjusted learning.

Recently, we began a long-term retention study and had to choose an acquisition criterion and retention intervals such that (1) retention performance would be different across the two retention intervals and not at the floor or ceiling for either and (2) the longer retention interval would not exceed an academic quarter. One of the most potent variables that emerged from examination of the data from nine previous experiments (Nelson, 1971, 1978; Nelson, Fehling, & Moore-Glascock, 1979) was the method of learning. The method involved standard list learning (in which the entire list is repeated during each study-test cycle until every item is recalled correctly on the same test cycle) vs. the method of adjusted learning (in which each correctly recalled item is dropped out of subsequent study-test cycles until every item is correctly recalled exactly once). The examination led to successful predictions of the retention levels in our study.

In this article, we report the results of two additional experiments along with the previous retention data so that other researchers might be in a better position to select long-term retention intervals and methods of learning.

EXPERIMENTS

For the nine published and two new experiments, the items were number-word pairs in which common words were randomly paired with two-digit numbers. Within each experiment, an attempt was made to keep ease of acquisition as homogeneous as possible across items (e.g., by using two-digit numbers from a narrow range of association values in the Battig and Spera, 1962, norms and by using only AA words from the Thorndike-Lorge, 1944, norms).

This project was partially supported by U.S. Public Health Service Grant MH-32205 to the first author. Requests for reprints should be sent to Thomas O. Nelson, Department of Psychology, University of Washington, Seattle, Washington 98195.

Method

Subjects. The subjects ($N > 20$ for every experiment) were undergraduates from Stanford University, the University of Illinois, or the University of Washington.

Procedure. During acquisition, recall from short-term memory was minimized either by blocking the list so that a minimum of 10 items intervened between the study and test of a given item (Nelson, 1971) or by interpolating 20-30 sec of number counting between study and test (Glanzer & Cunitz, 1966). The eventual long-term retention test was either subject-paced or paced at a rate of 10 sec/item. Across experiments, list length ranged from 16 to 24 pairs, study rate ranged from 4 to 12 sec/pair, and test rate ranged from 6 to 12 sec/pair; the response words were sometimes verbs, sometimes adjectives, and sometimes nouns.

For 7 of the 11 experiments, the method of adjusted learning (Battig, 1965; Gillette, 1936) was used with a criterion of one correct response per item. The remaining four experiments employed the standard study-test list learning method with a criterion of one completely correct trial on the entire list. The retention interval ranged from 1 to 7 weeks across experiments.

Results and Discussion

In spite of the procedural differences mentioned above, fairly systematic retention results were obtained by partitioning the 11 experiments in terms of retention interval and method of learning. Figure 1 shows the unweighted means of the percentages of items correctly recalled for experiments conducted at a given retention interval and with a given method of learning. These means were derived from the following mean percentages from each experiment. For the method of adjusted learning, the mean percentages are: after 1 week, 41% (new data); after 3 weeks, 28% (new data); after 4 weeks, 18% (Nelson, 1978, Experiment 2, Experiment 3; Nelson et al., 1979, Experiment 6); after 5 weeks, 14% (Nelson et al., 1979, Experiment 3); and after 7 weeks, 14% (Nelson et al., 1979, Experiment 7). For the method of standard list learning, the mean percentages are: after 1 week, 86% (Nelson, 1971, Experiment 1); after 2 weeks, 70% (Nelson, 1971, Experiment 1); and after 4 weeks, 52% (Nelson, 1978, Experiment 1) and 24% (Nelson et al., 1979, Experi-

Figure 1. Mean percentage of correct recall for experiments conducted at a given retention interval and with a given method of learning. The X data points show performance after the method of standard list learning, and the filled-circle data points show performance after the method of adjusted learning. Solid curves show the best-fitting exponential functions, and dashed curves show the best-fitting power functions (see text).

ment 1). As Figure 1 shows, the length of the retention interval has a potent effect regardless of the method of learning. The data are fit quite well by both exponential and power functions (cf. Wickelgren, 1972). The coefficient of determination (r^2) is .99 for the exponential function fit to the standard list learning method, .87 for the exponential function fit to the adjusted learning method, .92 for the power function fit to the standard list learning method, and .91 for the power function fit to the adjusted learning method. (There are not enough data to choose between the power and exponential functions, so both are shown in Figure 1.)

An important methodological finding is that for retention intervals from 1 week to at least 4 weeks, the standard list learning method yields much greater retention than the adjusted learning method. Performance at all retention intervals in this range is more than 100% greater after standard list learning than after adjusted learning. In terms of absolute differences, however, the number of extra items recalled after standard list learning (compared with adjusted learning) decreases as the retention interval becomes longer, indicating that the choice between the methods of learning is more crucial at shorter retention intervals.

Which method should be chosen depends on a number of factors: (1) If the goal is to have forgetting occur as quickly as possible, then the adjusted learning method is preferable. (2) If the goal is to obtain differential retention across two or more retention intervals, then additional criteria need to be considered. For example, considering the amount of negative deceleration in the functions shown in Figure 1, retention performance of 85% correct vs. 70% correct may be obtained fairly

easily by employing the standard list learning method with retention intervals of 1 and 2 weeks; however, the same retention values might be difficult to obtain accurately with the adjusted learning method, because retention drops below these values so quickly that there is less margin for any parameter-estimation error. By contrast, if the two retention intervals to be employed must be short (e.g., both intervals shorter than 1 week), then the adjusted learning method may be preferable because there is likely to be little difference in retention across short intervals after standard list learning. Other goals will lead to other criteria for choosing a method, but Figure 1 clearly indicates that the pattern of retention performance can be very different for the standard list learning and adjusted learning methods.

A general remark concerning differential retention for standard list learning vs. adjusted learning should be noted in addition to the parametric data presented above. The retention curve following standard list learning can be thought of as a weighted-average composite of the retention curves for items having one correct acquisition response, items having two correct acquisition responses, and so on (the weight for each retention curve being the number of items having the given number of correct acquisition responses). Figure 2 shows the retention data (collapsed over subjects and items) from the standard list learning

Figure 2. Percentage of correct recall as a function of retention interval (abscissa) and number of correct acquisition responses (curve parameter) for the standard list learning method. Subscripts indicate the number of correct acquisition responses (one vs. two vs. three vs. four or more), with data points shown as Xs. Solid curves show the best-fitting exponential functions.

method, with the curve parameter being the number of correct acquisition responses. Retention is better for items that are overlearned beyond one correct acquisition response (cf. Krueger, 1929; Postman, 1962); at shorter retention intervals, performance is also better following more than two correct acquisition responses than following exactly two correct acquisition responses, but this difference tends to disappear at longer retention intervals.¹ In the (probably hypothetical) extreme case in which every item is acquired on exactly the same trial, the retention curves for the standard list learning method vs. the adjusted learning method will be identical. Conversely, as the number of correct acquisition responses becomes more heterogeneous across items, the standard list learning method will produce overlearning of some items, which in turn will yield higher retention performance than the adjusted learning method.

The highest retention level for the standard list learning method occurs not when the number of correct acquisition responses is distributed rectangularly across items (e.g., when a 15-item list has 5 items with one correct acquisition response, 5 items with two correct acquisition responses, and 5 items with three correct acquisition responses), but when the number of correct acquisition responses for the hardest item is maximally different from the average number of correct acquisition responses for the remaining items (e.g., when a 15-item list has 1 item with one correct acquisition response and 14 items with three correct acquisition responses). Thus, an important consideration for the magnitude of the retention difference between these two methods is the heterogeneity across items in the number of trials required to learn each item; this is particularly important regarding the most difficult item because the extra trials required to learn this item also serve as overlearning trials for the remaining items in the standard list learning method. The method of adjusted learning helps to minimize overlearning and therefore is less vulnerable than the standard list learning method to heterogeneity among items.

In the experiments reported here, an attempt was made to employ homogeneous items. Nevertheless, different numbers of trials were required for different items. Moreover, it seems unlikely that memory researchers in general will be using lists (of any substantial length) in which all of the items will be learned on exactly the same trial. Therefore, differences can be

expected in long-term retention as a function of standard list learning vs. adjusted learning, and these differences can be quite large.

REFERENCES

- BATTIG, W. F. Procedural problems in paired-associate learning research. *Psychonomic Monograph Supplements*, 1965, 1(1, Whole No. 1).
- BATTIG, W. F., & SPERA, A. Rated association values of numbers from 0-100. *Journal of Verbal Learning and Verbal Behavior*, 1962, 1, 200-202.
- GILLETTE, A. L. Learning and retention: A comparison of three experimental procedures. *Archives of Psychology*, 1936, 28, 1-56.
- GLANZER, M., & CUNITZ, A. R. Two storage mechanisms in free recall. *Journal of Verbal Learning and Verbal Behavior*, 1966, 5, 351-360.
- KRUEGER, W. C. F. The effect of overlearning on retention. *Journal of Experimental Psychology*, 1929, 12, 71-78.
- NELSON, T. O. Savings and forgetting from long-term memory. *Journal of Verbal Learning and Verbal Behavior*, 1971, 10, 568-576.
- NELSON, T. O. Detecting small amounts of information in memory: Savings for nonrecognized items. *Journal of Experimental Psychology: Human Learning and Memory*, 1978, 4, 453-468.
- NELSON, T. O., FEHLING, M. R., & MOORE-GLASCOCK, J. The nature of semantic savings for items forgotten from long-term memory. *Journal of Experimental Psychology: General*, 1979, 108, 225-250.
- POSTMAN, L. Retention as a function of degree of overlearning. *Science*, 1962, 135, 666-667.
- THORNDIKE, E. L., & LORGE, I. *The teacher's wordbook of 30,000 words*. New York: Columbia University Press, 1944.
- WICKELGREN, W. A. Trace resistance and the decay of long-term memory. *Journal of Mathematical Psychology*, 1972, 9, 418-455.

NOTE

1. Note that the retention differences shown in Figure 2 are probably greater than would be obtained if items were randomly assigned to different numbers of correct acquisition responses. For example, the one-correct-acquisition-response items here would, by definition, tend to be idiosyncratically more difficult items than would the two-correct-acquisition-response items because the former were not correctly recalled until the final acquisition trial, whereas the latter were correctly recalled at least one trial prior to the final acquisition trial. Thus, the differences in the retention curves shown in Figure 2 represent the composite effects of number of correct acquisition responses and idiosyncratic item difficulty. These two effects go hand in hand when the standard list learning method is employed.

(Received for publication July 27, 1979;
revision accepted December 19, 1979.)