INSTRUMENTATION & TECHNIQUES

Postcriterion overlearning reduces the effectiveness of the method of adjusted learning

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

One goal of the method of adjusted learning ("dropout procedure") is to equate the strength of different items for subsequent retention. Some studies have yielded results in accord with this goal while others have not, with no known reason for the discrepancy. We manipulated the number of postcriterion overlearning trials and found that the adjusted learning procedure was effective in equating items for subsequent retention when learning terminated after one correct response per item but ineffective when postcriterion overlearning trials occurred. This finding is especially important for those versions of the adjusted learning procedure, such as the frequently utilized variant developed by Battig (1965), that incorporate overlearning of the to-be-equated items.

The method of adjusted learning ("dropout procedure") is often used to equate the degree of learning among items so that they will have equivalent subsequent retention. The procedure consists of alternating study and test phases in which the individual items are dropped out of the list when they have reached some specified criterion of learning (e.g., one correct recall per item). The study-test sequence is repeated with the shortened list until all items eventually reach criterion, at which point the learning session terminates.

This technique was first employed by Woodworth (1914). In almost every decade since the 1930s, variants of this method have been devised (e.g., Battig, 1965; Cavanaugh & Parkman, 1971; Gillette, 1936; Madden, Adams, & Spence, 1950; Nelson, Shimamura, & Leonesio, 1980).

The most highly cited—and perhaps most frequently utilized—variant of the method of adjusted learning is the one developed by Battig (1965). In that version, alternating study and test trials occur (with correctly recalled items being dropped out of the list) until all items have been dropped out, which defines the end of the first cycle. Then the second cycle of study-test trials occurs, with all of the items back in the list and correctly recalled items dropped out as in the first cycle. These cycles continue until some specified number of items is correctly recalled on the first test trial of a given cycle (e.g., six of eight), at which point

learning terminates. Notice that this procedure necessarily involves overlearning of the items beyond the first correct recall (at the very least, overlearning from the second recall of each item, and perhaps even more overlearning if more than two cycles occur before the subject reaches the overall-recall termination criterion). Subsequently, a retention test occurs, and the hope is that the dropout procedure will have produced equivalent learning of all the items such that there will be no relationship between the speed of learning for the various items and their subsequent retention.

Some investigators have found that the method of adjusted learning is effective for producing homogeneous subsequent retention across items (e.g., Cavanaugh & Parkman, 1971). However, others report heterogeneous retention in which the quickly learned items are more likely to be recalled than are the slowly learned items on a subsequent retention test (e.g., Montague & Kiess, 1968).¹ The reason for these different findings is currently unknown.

Underwood (1964) highlighted the potential theoretical importance of postcriterion overlearning for producing subsequent heterogeneity of retention for items that are learned to the same criterion. He suggested that an overlearning trial would benefit retention more for a quickly learned item than for a slowly learned item. However, we are not aware of any direct empirical examination of the effect of postcriterion overlearning on subsequent heterogeneity/homogeneity of retention for different items. Consequently, we report data from research in which the number of overlearning trials was manipulated, and the effectiveness of the dropout procedure for producing homogeneous retention of different items was assessed.

This research was supported by Grant MH-32205. Reprint requests should be sent to Thomas O. Nelson, Department of Psychology, University of Washington, Seattle, Washington 98195.

METHOD

The present findings emerged in the context of a larger study described in detail elsewhere (Nelson, Leonesio, Shimamura, Landwehr, & Narens, 1982). The independent variable of interest in this study consisted of repeated measures on the within-subjects variable of one vs. two vs. four correct responses per item during acquisition. The subjects were 50 University of Washington undergraduates whose participation partially fulfilled a course requirement. The items were 18 number-noun pairs of the form 48-dollar. The 18 stimulus terms were two-digit numbers, and the 18 response terms were common nouns. All stimuli were presented via an Apple II microcomputer and displayed on a television monitor. The computer-controlled presentation of the items made the method of adjusted learning especially easy to implement.

Each number-noun pair was randomly assigned to a learning criterion of one, two, or four correct responses. There were six pairs at each of the three criterion levels. During study, pairs were shown individually at a 16-sec rate. The subject read aloud each number-word pair as it was presented and then remained silent until the next pair appeared. Following study, 26 sec of backward number-counting occurred to minimize recall from short-term memory. Then the test phase began. Each number stimulus was presented, and the subject attempted to recall the correct noun response. The subject entered his/her response on the computer keyboard. The order of presentation for the items was random for both the study and test phases. After each study-test sequence, all pairs that reached their preassigned criterion of correct responses were removed (e.g., a one-correct item was removed after being recalled correctly on one test trial; a two-correct item was removed after being recalled correctly on two test trials). Thus, the next study-test sequence included only those items that had not yet reached criterion during the previous study-test sequence. The study-test sequences continued until every item was learned to its preassigned criterion. Without being informed in advance of the impending retention test, the subjects returned 4 weeks later and received a recall test for each of the 18 pairs, using the same procedure as in the test phase of acquisition.

RESULTS

Before describing the primary results, two peripheral findings should be mentioned. First, once an item was correctly recalled during acquisition, it was almost always correctly recalled on subsequent acquisition trials. That is, the mean probability of correct subsequent recall during acquisition, after the first correct recall of that item, was .96. This implies that when the present procedure is employed (in which some filler activity is interpolated between study and test to minimize the role of recall from short-term memory), one can speak almost interchangeably about different degrees of learning (as defined by different criteria such as one, two, or four correct recalls) or about postcriterion overlearning trials.

Second, the independent variable of overlearning trials had a reliable effect on subsequent recall retention [$F(2) = 18.99$, $p < .05$]. Sign tests showed that the probability of recall retention was reliably greater for items recalled four times during acquisition (.393) than for those recalled two times (.287). Probability of recall retention for those recalled two times

was, in turn, reliably greater than for items that had been recalled one time during acquisition (.182). This is in accord with the standard finding that overlearning produces increases in subsequent retention.

Turning to the main measure of interest, for each subject, we computed across items the Goodman-Kruskal gamma correlation between the number of acquisition trials to the first correct response (i.e., item difficulty) and recall on the retention test. When this correlation was computed over all 18 items for a given subject, the mean gamma (for all subjects) between retention and the number of acquisition trials to the first correct response was a reliable $-.20$, indicating that retention decreases with increases in the number of trials required for the first correct response. That is, the harder items to learn were also the harder items to retain. Hence, the adjusted learning procedure was not successful at equating acquisition for items of different degrees of difficulty.

Next, and of primary importance, we teased apart the previously mentioned overall ineffectiveness of the adjusted learning procedure by computing the same gamma correlation (between retention and the number of acquisition trials to the first correct response) separately for the six items that had been correct once, twice, and four times during acquisition. The three gammas were computed individually for each subject's data. The mean gamma for one-correct items was .02, for two-correct items was $-.15$, and for four-correct items was $-.25$. Only the last of the three mean gammas is reliably different from zero ($p < .05$). Thus, the degree of overlearning after one correct recall appears to modulate the relationship between item difficulty and subsequent retention, such that extra overlearning trials tend to reduce the effectiveness of the dropout procedure for equating different items for subsequent retention.

DISCUSSION

Item difficulty appears to have a negligible effect on subsequent retention for items that are acquired to a low degree of learning (e.g., one correct recall during acquisition) but has reliable effects at higher degrees of learning (e.g., four correct recalls during acquisition). Apparently, overlearning trials can amplify the effect of heterogeneous item difficulty, such that otherwise undetectable differences become detectable.²

This finding is in accord with Underwood's (1964) theoretical analysis that overlearning trials should be more beneficial for easy items than for difficult items, thereby accentuating the differences among items. However, this result disconfirms Battig's notion that differences among items "become substantially reduced and may be virtually eliminated through the use of more stringent learning criteria" (1965, p. 6). To further the theoretical interpretation of the effects of overlearning

on the relationship between item difficulty and subsequent retention, future research should examine this same kind of effect for different percentages of overlearning (cf. Postman, 1962), as well as for different fixed numbers of overlearning trials. Whether a constant percentage of overlearning (as contrasted with a constant number of overlearning trials) will enhance the effects of item difficulty on subsequent retention is currently unknown.

The methodological ramifications of the present findings are straightforward. A fixed number of overlearning trials beyond one correct recall tends to reduce the effectiveness of the adjusted learning method for equating subsequent retention for items of different levels of acquisition difficulty. This is especially important for Battig's (1965) variant of the method of adjusted learning because, in his method, items that have already been correctly recalled continue to receive overlearning trials until the entire list is recalled to some overall criterion. Hence, extra overlearning trials are the rule rather than the exception for Battig's variant; this may help to explain why different investigators have found different results about the effectiveness of the adjusted learning procedure for equating items for subsequent retention. Moreover, some researchers using Battig's variant do not report the number of cycles before the subject attains the overall-list criterion (e.g., Montague & Kiess, 1966; Wearing & Montague, 1970). In these studies, therefore, the reader cannot know the number of overlearning trials after the first correct recall of a given item. Because overlearning modulates the effectiveness of the adjusted learning procedure, future studies attempting to equate different items for subsequent retention should give careful consideration to the number of postcriterion overlearning trials that occur during acquisition.

REFERENCES

- BATTIG, W. F. Procedural problems in paired-associate learning research. *Psychonomic Science Monograph Supplements*, 1965, 1(Whole No. 1), 1-12.
- CAVANAUGH, P. J., & PARKMAN, J. A. A comparison of the study-recall and anticipation methods in steady state paired-associate learning. *Psychonomic Science*, 1971, 22, 361-363.
- GILLETTE, A. L. Learning and retention: A comparison of three experimental procedures. *Archives of Psychology*, 1936, 28.
- MADDEN, M. S., ADAMS, J. A., & SPENCE, S. A. Memory-drum vs. adjusted-learning techniques in the study of associative interference in learning by paired associates. *American Journal of Psychology*, 1950, 63, 186-195.
- MONTAGUE, W. E., & KIESS, H. O. Test of a procedure to control inter-pair learning and the effect of repeated, post-criterion test trials on retention. *Psychonomic Science*, 1966, 6, 191-192.
- MONTAGUE, W. E., & KIESS, H. O. The associability of CVC pairs. *American Journal of Psychology*, 1968, 78, 1-23.
- NELSON, T. O., LEONESIO, R. J., SHIMAMURA, A. P., LANDWEHR, R. F., & NARENS, L. Overlearning and the feeling of knowing. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 1982, 8, 279-288.
- NELSON, T. O., SHIMAMURA, A. P., & LEONESIO, R. J. Large effects on long-term retention after standard list learning vs. adjusted learning. *Behavior Research Methods & Instrumentation*, 1980, 12, 42-44.
- POSTMAN, L. Retention as a function of degree of overlearning. *Science*, 1962, 135, 666-667.
- UNDERWOOD, B. J. Speed of learning and amount retained: A consideration of methodology. *Psychological Bulletin*, 1954, 51, 276-282.
- UNDERWOOD, B. J. Degree of learning and the measurement of forgetting. *Journal of Verbal Learning and Verbal Behavior*, 1964, 3, 112-129.
- WEARING, A. J., & MONTAGUE, W. E. A test of the Battig procedure for controlling the level of individual item learning in paired-associate lists. *Behavior Research Methods & Instrumentation*, 1970, 2, 9-10.
- WOODWORTH, R. S. A contribution to the question of "quick learning, quick forgetting." *Psychological Bulletin*, 1914, 11, 58-59.

NOTES

1. Wearing and Montague (1970) reported that items with fewer errors had shorter latencies. However, if slow responders made more errors than fast responders, the former could have been overrepresented at the higher error levels, resulting in the observed increase of latency with item difficulty. Cavanaugh and Parkman (1971), avoiding this confound, found no correlation between number of errors and recall latency. Our own unpublished data also show no such correlation.

2. It is possible that these arguments concerning heterogeneous items can be extended to heterogeneous learners (cf. Underwood, 1954).

(Received for publication July 28, 1981;
revision accepted February 8, 1982.)