*FACILITATION OF THE ACQUISITION AND RETENTION OF
SIGHT-WORD VOCABULARY THROUGH TOKEN
REINFORCEMENT¹*

BENJAMIN B. LAHEY AND RONALD S. DRABMAN

FLORIDA TECHNOLOGICAL UNIVERSITY

Two groups of second-grade children were taught to read orally 30 words and were later tested for retention. One group of subjects received verbal consequences, while the other group received verbal consequences plus token reinforcement. The token group reached criterion in significantly fewer trials and correctly read significantly more words on the retention tests after the longer retention intervals.

Reinforcement principles have been applied to the problem of teaching reading in a number of basic and applied studies over the past 10 yr (Corey and Shamow, 1972; Gray, Baker, and Stancyk, 1969; Staats and Butterfield, 1965; Staats, Finley, Minke, and Wolf, 1964; Staats, Minke, Finley, Wolf, and Brooks, 1964; Staats, Staats, Shutz, and Wolf, 1962; Whitlock and Bushell, 1967). Perhaps the most encouraging of the applied studies is the report of the Englemann-Becker program used in Project Follow Through. Their program has been shown to produce increases of over one grade level in reading achievement per academic year in low-income groups that typically show an average gain of about one-half grade level (Becker, *unpublished*). In spite of the considerable amount of research on reading, however, several basic factors remain uninvestigated. For example, little attention has been paid to the fundamental

question of the effects of reinforcement on reading retention.

In fact, this question has received little attention in the experimental human learning and retention literature. Harley (1965*a*, 1965*b*) and Weiner and Walker (1966) looked at the effects of response consequences given during retention tests, with equivocal results. But few investigations have looked at the effects on later retention of reinforcement given during acquisition.

An exception is a study by Risley and Wolf (1967). In this study, a nonverbal child was taught to label orally a series of pictures through modelling and reinforcement. When he reached a criterion of giving the unprompted correct label to a given picture on its first presentation on three successive days, that picture was dropped from training. The child was tested for retention by presenting the picture again after the child had reached criterion on 10 succeeding pictures. Correct labelling responses to one set of pictures were followed by verbal praise; responses to another set by praise plus ice cream. Analysis of the acquisition data showed that praise plus ice cream resulted in fewer trials to criterion, but that there was no significant difference in the per cent of labels retained under the two conditions.

In view of the importance of this area of research to the development of an effective technology of education, it would appear that more

¹This study was partially supported by grant 1-RO3-MH21948-01 from the National Institute of Mental Health to the first author and grant 1-RO3-MH24502-01 from the National Institute of Mental Health to the second author. Reprints may be obtained from either Ben Lahey, now at the Department of Psychology, University of South Carolina, Columbia, South Carolina 29208, or Ron Drabman, Department of Psychology, Florida Technological University, Box 25000, Orlando, Florida 32816. Our thanks to Bill Pearce, Maryanne Cook, Kirsten Comish, and Bill Hunt for their help in running subjects, and to the people of the Seminole County School System for their cooperation and assistance.

research would be justified. The present investigation examined the effects of token reinforcement on acquisition and retention of a reading behavior.

METHOD

Subjects

Sixteen second-grade students, from a small town in a predominantly rural county, were drawn from the middle of three second-grade reading groups. Two of the children were black, one was Cuban, and the remainder were white.

Setting

Subjects were brought from their classrooms to the experimental room four at a time. The experimental room contained four tables separated from one another by opaque screens. Each subject sat facing an experimenter across a table. The presence of four subjects in the room created a background of low, indistinct noise. Data were recorded by the experimenter behind a low opaque screen. The subjects returned to their rooms after each session as a group. Two male and two female college student experimenters were used. Each experimenter worked with two subjects in each group at different times.

Materials

Stimulus materials were drawn from the "Basic Sight Vocabulary Cards" (Dolch, 1949). Each 2 by 3.5 in. (5 by 8.9 cm) card contained a single second-grade vocabulary word.

Procedure

Selection of stimulus items. Each subject participated for five twice-weekly sessions. During the first session, each subject was shown a large set of the sight-word cards one at a time in a random sequence and asked to read them. Each card that was not read correctly was selected as a stimulus item for that subject, up to a total of 30 words. Since the response definition was unambiguous, no reliability checks were made,

as is typical in investigations of this sort (Corey and Shamow, 1972).

Acquisition. The subjects were randomly divided, with sex balanced, into two equal groups that were given different training procedures. Both groups were trained to read one set of 10 sight cards during each of the three succeeding sessions. Subjects in both groups were trained to a criterion of one trial without errors, or until the session reached 1 hr in length. A trial was defined as one presentation of the list of 10 words. All 10 words in the list were presented on every acquisition trial. Only one subject in the No-Token group failed to learn all of the words within 1 hr.

On the first trial of each of the three acquisition sessions, each subject in both groups was shown 10 of the words in his list, one at a time. Approximately 1 sec after the experimenter held the word up, he pronounced it to the subject. On each successive trial, the experimenter asked the subject to try to read each word out loud.

In the Token group, subjects were told before the first acquisition session that they would be shown some words and that the experimenter would help them learn to read them. The subjects were told that they would receive a token every time they read a word correctly and that the tokens could be exchanged later for pennies. The subjects were not told the exchange rate of tokens for pennies until the first exchange, which took place after the first acquisition session.

On each trial after the first, the subject was shown each of the 10 cards individually and asked to read them. If the response was correct, the child was told, "Right", and a token was dropped into a dish in front of him. If the response was incorrect, the child was told, "No, its —", and was told the correct answer. Tokens were not given for incorrect responses. This procedure was repeated several times through the set of 10 words until the subject correctly read all 10 words in the list one time through. Three such acquisition sessions were carried out, with the subjects learning a total of 30 sight words. At the end of each session, the tokens

were exchanged at the rate of five tokens per penny.

The procedure in the No-Token group was the same as that in the Token group except that tokens were not given for correct responses. Subjects in this group were told in the instructions that they would get some pennies at the end of the experiment. They were given the same number as the total earned by randomly assigned yoked partner in the Token group.

Retention tests. Immediately after the last trial of the third acquisition session, each subject in both groups was shown all 30 words one at a time in random order and asked to read them. The subject in the No-Token group that did not reach criterion during acquisition was not given the retention tests. No tokens or verbal response consequences were given. All 30 words were presented, 10 of which had been learned the same day, 10 of which had been learned five days earlier, and 10 of which had been learned seven days before the test. Two days later, the cards were reshuffled and presented again in the same manner for a second retention

test. On this presentation, the three sets of 10 words had been learned two, seven, and nine days before the test.

RESULTS

The average number of trials to criterion in learning each set of 10 sight words in each group and the number of words correctly read on the retention tests are shown in Table 1, and the number of correct responses per trial is shown for each subject in Figure 1. On the average, subjects in the No-Token group took twice as many trials to reach criterion on the three sets of words as did the Token group. The difference in the total number of trials required to reach criterion on the three sets of sight words was significant at the 0.05 level ($t = 2.30$, 14df, two tailed). An analysis of variance showed that there were no significant effects of experimenters ($F(3,8) = 5.70$) or savings in learning successive sets of words ($F(2,16) = 1.19$). No interactions approached significance. The individual graphs in Figure 1 show that the measure of trials to

Table 1

Number of trials to criterion in each acquisition session and number of words recalled on each retention test.

Tokens					No Tokens				
Trials to Criterion per Acquisition Session			Words Recalled per Retention Test		Trials to Criterion per Acquisition Session			Words Recalled per Retention Test	
I	II	III	I	II	I	II	III	I	II
2	2	1	25	27	3	2	1	23	22
2	2	5	25	26	5	3	2	21	21
3	2	2	25	28	5	6	10	23	19
3	3	3	25	25	9	6	11	12	19
5	2	4	20	17	10	7	12	16	20
7	14	5	25	27	13	15	6	16	20
9	9	3	14	20	16	15	15	—	—
10	5	5	22	17	17	11	9	14	14
$\bar{X} = 5.1$	4.9	3.5	22.6	23.4	9.8	8.1	8.2	17.9	19.3

Note: Subjects are rank ordered on the basis of the number of trials to criterion on the first acquisition session. One set of 10 sight words was learned in each of the three acquisition sessions to a criterion of one trial (one presentation of each of the 10 words in the set) without errors. All three sets of words were presented once during each retention test for a maximum of 30 correct responses possible on each of the two tests. The mean of each column is presented in the bottom row.

Fig. 1. The number of words correctly read on each of the acquisition trials until criterion was reached of one trial through the list of 10 words without errors. Each line represents the performance of one subject. Graphs are separated for the three sets of 10 words each learned during the three acquisition sessions and by experimental groups.

criterion is indicative of overall performance. Subjects with larger numbers of trials to criterion generally made fewer correct responses per trial than those with fewer trials to criterion.

The total number of words read correctly on each retention test is also presented in Table 1. Subjects in the Token group recalled more words on the first retention test ($t = 2.19$, 13df, two tailed, $p < 0.05$) and on the second retention test ($t = 2.14$, 13df, two tailed, $p < 0.06$). An analysis of variance of the combined total number of words recalled on both retention tests added together also showed the effect of tokens to be significant ($F(1,7) = 6.99$, $p < 0.05$), as well as the effects of retention interval (the number of days between when a list was learned and the retention test) ($F(2,14) = 11.15$, $p < 0.01$). The effect of experimenters was not significant ($F(3,7) = 1.01$).

The interaction between tokens and retention interval was also significant ($F(2,14) = 7.07$, $p < 0.01$). This reflects the fact that differences in retention between the two groups increased as the retention interval increased. As shown in

Figure 2, both groups recalled approximately the same number of words at the shortest interval, but the Token group recalled more words

Fig. 2. Average number of sight words read correctly per 10-word list in the combined retention tests. The shortest retention interval was zero days on the first test and two days on the second; the medium retention interval was five days on the first test and seven days on the second; and the longest retention interval was seven days on the first test and nine days on the second test.

at the longer intervals. No other interactions approached significance.

DISCUSSION

Although the question of retention is of fundamental importance to any approach to education, few behavioral investigations of instruction have previously looked at this topic. Fortunately, from the point of view of the development of a technology of education, the present results suggest that the same procedures that facilitate acquisition also facilitate retention. The significant interaction between tokens and retention interval is particularly important in this respect, as it suggests that the importance of token reinforcement increases as the length of the retention interval increases.

Because each subject was taught a unique set of words (the words the subject missed during the screening session), it is possible that the Token group learned an "easier" set of words than did the No-Token group. The random assignment of subjects to groups should have reduced the likelihood of this possibility, but an additional comment can be made. Because the words for each subject were drawn from a fixed set of 200 words, there was considerable overlap in the words used by each group (77% of the words were used in both groups). On the other hand, because each subject was taught only 30 of these words, there was not much overlap in the words used within each experimental group. The average number of times that a word was used in the Token group was 1.9 and was 1.8 times in the No-Token group. The fact that there was considerable overlap in the populations of words used by each group, but only a small amount of overlap among subjects within each group, suggests that the differences between groups cannot be reasonably attributable to differences in the words learned.

The present results on acquisition of reading are similar to those obtained by Risley and Wolf (1967), who found that a nonverbal child learned to label pictures in fewer trials when

praise and ice cream followed correct responses than when just praise followed correct responses. However, the present results on retention are different, in that Risley and Wolf found no differences in the percentage of labels retained. But, the many differences between the present study and the one by Risley and Wolf, including the different types of subject populations employed, the somewhat different tasks and consequences used, and the different criteria for acquisition, make it difficult to draw direct comparisons.

It is also interesting to compare the present results with those of Corey and Shamow (1972). In their study, children were taught to read sight words by pairing them with pictures of the object labelled by the word. A comparison was made of methods in which pictures were removed abruptly or gradually faded out. The fading technique produced both faster acquisition and better retention. Since both groups in the study by Corey and Shamow (1972) received the same response consequences, further investigations may reveal that any procedure that facilitates acquisition will also facilitate retention.

A comment should also be made as to the importance of the reading behavior used in this study. Most modern approaches to reading attempt to teach children in a way that they can pronounce any word, whether they have seen it before or not (*e.g.*, "phonics system"). All reading systems, however, must teach the child a large number of irregular words. These words must be learned as individual units. Teaching them as sight words might prove to be an effective component of an overall instructional system.

REFERENCES

- Becker, W. *Some effects of direct instruction methods in teaching disadvantaged children in Project Follow Through*. Unpublished paper presented at the International Symposium on Behavior Therapy, Minneapolis, Minnesota, 1972.
- Corey, J. R. and Shamow, J. The effects of fading on the acquisition and retention of oral reading. *Journal of Applied Behavior Analysis*, 1972, 5, 311-315.

- Dolch, E. W. *Basic sight vocabulary cards*. Champaign: Garrard, 1949.
- Gray, B. B., Baker, R. D., and Stancyk, S. E. Performance determined instruction for training in remedial reading. *Journal of Applied Behavior Analysis*, 1969, 2, 255-263.
- Harley, W. F. The effect of monetary incentive in paired associate learning using a differential method. *Psychonomic Science*, 1965, 2, 377-378.
(a)
- Harley, W. F. The effect of monetary incentive in paired associate learning using an absolute method. *Psychonomic Science*, 1965, 3, 141-142.
(b)
- Risley, T. R. and Wolf, M. M. Experimental manipulation of autistic behaviors and generalization into the home. *Behaviour Research and Therapy*, 1967, 5, 73-88.
- Staats, A. W. and Butterfield, W. H. Treatment of non-reading in a culturally deprived juvenile delinquent: an application of reinforcement principles. *Child Development*, 1965, 3, 925-942.
- Staats, A. W., Finley, J. R., Minke, K. A., and Wolf, M. Reinforcement variables in the control of unit reading responses. *Journal of the Experimental Analysis of Behavior*, 1964, 1, 139-149.
- Staats, A. W., Minke, K. A., Finley, J. R., Wolf, M., and Brooks, L. O. A reinforcer system and experimental procedure for the laboratory study of reading acquisition. *Child Development*, 1964, 35, 209-231.
- Staats, A. W., Staats, C. K., Schutz, R. E., and Wolf, M. The conditioning of textual responses using "extrinsic" reinforcers. *Journal of the Experimental Analysis of Behavior*, 1962, 5, 33-40.
- Weiner, B. and Walker, E. L. Motivational factors in short-term retention. *Journal of Experimental Psychology*, 1966, 71, 190-193.
- Whitlock, C. and Bushell, D. Some effects of "back up" reinforcers on reading behavior. *Journal of Experimental Child Psychology*, 1967, 5, 50-57.

Received 16 July 1973.

(Revision requested 8 October 1973.)

(Revision requested 10 January 1974.)

(Final acceptance 18 February 1974.)