MODIFICATION OF DEFICITS IN READING FOR COMPREHENSION

BENJAMIN B. LAHEY¹, MICHAEL P. MCNEES², AND CATHY C. BROWN

FLORIDA TECHNOLOGICAL UNIVERSITY

Subjects were asked to read a passage orally and then to answer questions about the passage. When praise and pennies were given for correct answers, the percentage of correct answers increased in two sixth-grade subjects whose reading for comprehension was tested to be 2 yr below grade level. The behavior of these subjects was compared to that of two subjects whose reading for comprehension was tested to be on grade level. Although no evidence for changes in the accuracy of answering comprehension questions is found in the previous literature, the percentage of correct answers in the children with deficits increased to approximately the same level as the children tested to be on grade level.

Over the past decade, a number of investigators have reported successful applications of operant techniques to the teaching of reading behaviors. Corey and Shamow (1972), Gray, Baker, and Stancyk (1969), Staats and Butterfield (1965), Staats, Minke, Goodwin, and Landeen (1967), Staats, Finley, Minke, and Wolf (1964), Staats, Staats, Schutz, and Wolf (1962), and Whitlock and Bushell (1967) found reinforcement procedures to be effective in producing significant changes in the rate and accuracy of oral reading of single words and passages.

Positive results have not been found for all reading behaviors, however. A primary purpose for reading instruction is to teach individuals to answer accurately questions about the passages they have read ("reading for comprehension"). Several studies have looked at this behavior. Staats and Butterfield (1965) and Staats *et al.* (1967) gave contingent tokens for accurate oral reading of words and passages and for correct answers to questions about silently read passages. Significant pre-posttest changes on standardized reading tests were found for a single subject in

Staats and Butterfield (1965), but no significant changes were found for 18 subjects in Staats *et al.* (1967). Negative results were also found by Camp and Van Doorninck (1971) who used similar procedures. Separate direct measures of the accuracy of answering comprehension questions and results for the comprehension subtests of the standardized reading tests were not presented in any of these studies, making it impossible to draw conclusions about the modifiability of reading for comprehension. These data suggest, however, that operant procedures would not be effective. Considering the social significance of this behavior, this represents an important gap in the educational behavior analysis literature.

The subjects in the present experiment came from a population that is interesting from both applied and theoretical perspectives. If reading is a complex of separate behaviors, it would not be surprising to find children who have acquired one or more reading behaviors less well than other reading behaviors due to variations in reading instruction and other conditions. It appears to be the case that many children do answer comprehension questions markedly less well than they read orally. This is a familiar pattern to many special education teachers and has been noted in the literature: "In addition

¹Reprints may be obtained from Ben Lahey, Department of Psychology, University of South Carolina, Columbia, S.C. 29208.

²Now at Johnny Cake Child Study Center, Mansfield, Arkansas.

there are those children who can apparently decode quite well but have little or no comprehension of the material". (Gray *et al.*, 1969, p. 262). This pattern may result in many cases from a tendency of teachers to focus on oral reading instruction, but this remains to be demonstrated. In any case, such children are a puzzle to their teachers as they read well (orally), but do poorly in subjects that require reading (for comprehension).

METHOD

Subjects

Two of the subjects were a black male and a white female from low income families who attended the sixth grade of a rural public school. They made poor grades and were selected on the basis of their pattern of performance on the Spache Diagnostic Reading Scales (1963). Separate oral reading and comprehension scores can be obtained with this scale.

The type of pattern sought was explained to the teacher and she selected several students from her class for testing. The two subjects selected both read on grade level in oral reading, and showed a large discrepancy between oral reading and comprehension. Subject A was 11 yr, 10 months at the start of the study. Her Spache oral reading grade level was 6.5 and her comprehension grade level was 4.5. Subject B was 11 yr, 3 months as the start of the study. His Spache oral reading grade level was 6.5 and his comprehension grade level was 4.5.

A goal of this study was to bring these children up to grade level on reading for comprehension. To assess the degree of improvement, their behavior was compared to that of two subjects whose reading for comprehension was tested to be on grade level. This strategy was chosen over a pre-posttest comparison on comprehension subtest scores because of the small number of subjects, the lack of information on test-retest improvements in scores in the absence of training ("practice effects") for such individual reading tests, and the previous indica-

tions of the low test-retest reliability of such tests with remedial reading populations (Staats *et al.*, 1967).

Due to the latter characteristic of standardized reading tests, this present method of assessment is not without problems, either. If such tests are not reliable, then it cannot be said with confidence that the scores that put the children at or below grade level during the selection procedure were accurate indications of their actual reading levels. This method was simply chosen as the better assessment strategy of the available methods. Fortunately, a degree of empirical validation of the selection procedure was possible in this study. Had the baseline performance of the two subjects tested to be below grade level in reading for comprehension not been consistently similar, as well as the higher performance of the two subjects tested to be on grade level, this assessment strategy could not have been employed.

The comparison subjects were tested to have both oral and comprehension grade levels of 6.5 on the Spache Diagnostic Reading Scales. Both subjects were low-income white males from the same class as Subjects A and B. Subject C was 11 yr, 7 months at the start of the study. Subject D was 11 yr, 1 month.

Setting

All phases of the procedure were carried out in an unused classroom with two subjects at a time (A and B, or C and D) seated side-by-side in school desks facing the male experimenter. The children were brought to the classroom two or three times weekly during the normal school day for 30-min sessions.

Procedure

Each child was given a set of several pages of printed reading material from the sixth-grade reader of *The Basic Reader* series (Robinson, Monroe, Huck, and Jenkins, 1965) on which short passages (five sentences followed by 10 paragraphs of two to five sentences) had been

circled in black. The set of material for each child was different but comparable. In alternation, each child was instructed to read aloud one of the encircled passages from the series. The child was then immediately asked from one to four simple factual questions about the passage. This continued in alternation until all of the passages had been read. During each session, each subject was asked from 27 to 35 questions. The questions were prepared in advance in the same form as questions in the comprehension sections of most standardized individual reading tests. The number of questions following each passage was determined by the content of the passage.

Only verbatim or approximate synonym answers were counted as correct. For example, a subject might be asked to read the following passage:

"The dog chased the cat up the tree. The cat ran as fast as it could." After reading the passage, the subject would be asked:

"The dog chased the cat up the what?" An answer of "tree" or answers of the form "the oak" or "the big bush" would be counted as correct. Any other type of answer would be counted as incorrect. Throughout the study, the same form of question based on passages from the same reader was used, but all of the passages were different within each session and all of the sets of passages were different for each session of the study. The dependent variable was expressed as the per cent of questions answered correctly.

Occasionally, the subjects were unable to read part of a passage correctly. In these instances, they were instructed to pronounce the word as best as they could. If such a word was involved in the answer, the experimenter accepted the same pronunciation given by the subject as being correct. Such instances were very rare and almost always involved proper nouns. The experimenter observed no occasions of subjects supplying one another with answers.

A second experimenter was present on eight occasions for the purpose of taking inter-

observer reliability checks. This experimenter was positioned so that he could not see the records made by the first experimenter. He was instructed to make independent judgements and was specially cautioned to judge the subject's responses before the first experimenter delivered the response consequences during the contingency phases and not to change any judgements after hearing the consequences. It is possible, however, that such judgements were not entirely independent during the contingency phases. On the other hand, the judgements of the second experimenter can be said to be independent during the two baseline phases when there were no response consequences. The close similarity of the reliability figures for the eight checks made during contingency phases and the eight checks made during baseline phases suggest that all of these reliability checks were independent. Reliability on an answer-by-answer basis ranged from 0.89 to 1.00 with a mean of 0.98. The sessions on which checks were taken are indicated in Figures 1 and 2.

Baseline I. The subjects were instructed to read each encircled passage one at a time and were immediately asked one or more questions following each passage. The experimenter said or did nothing contingent upon their answers. The children alternated until all of the passages in each of their sets of materials had been read. The number of sessions during each phase is indicated in Figure 1 for Subjects A and B and Figure 2 for Subjects C and D.

Pennies and Praise I. The procedure during this phase was the same as in Baseline I, except that social praise (e.g., "right") and a penny were made contingent upon each correct answer. Following incorrect answers, the experimenter said nothing and did not give the subject a penny.

Baseline II. Following the first Praise and Pennies phase, a second baseline was carried out with conditions identical to those in Baseline I.

Pennies and Praise II. Following Baseline II, the conditions of the first Praise and Pennies phase were reintroduced.

RESULTS

The results for Subjects A and B are presented in Figure 1. The introduction of the praise-and-pennies contingency improved the percentage of questions answered correctly. Performance during this phase stabilized at a high level, with both subjects answering over 90% of the questions during the last five sessions. Percentage of correct answers fell again during Baseline II to a level within the upper range of Baseline I, but rose again to a level at or above 90% when the contingency was reintroduced.

The results for Subjects C and D are shown in Figure 2. The introduction and removal of the pennies-and-praise contingencies had no noticeable effect on the behavior. Across sessions, there appeared to be a slight increase in the percentage of correct answers and a slight reduction in variability, but there was no systematic change in behavior associated with the introduction and removal of the contingency.

The behavior of these subjects might be taken as the "ceiling" performance of this behavior—the highest level of performance on reading for comprehension under the conditions of the contingency for subjects from this student population. In this light, it is interesting to compare the performance of both sets of subjects. During Baselines I and II, the performance of Subjects A and B, who were below grade level on the reading for comprehension test, was clearly lower than the performance of Subjects C and D, who were at grade level during the reading comprehension test. During Pennies and Praise I and II, however, the percentage of correct answers of Subjects A and B rose into the "ceiling" level of performance of Subjects C and D.

DISCUSSION

No clear conclusions can be drawn from the previous literature as to the modifiability of read-

Fig. 1. Per cent of questions answered correctly by Subjects A and B, whose oral reading was tested to be on grade level and whose reading for comprehension was tested to be below grade level. Subject A drawn in dotted line, Subject B in solid line.

Fig. 2. Per cent of questions answered correctly by Subjects C and D, whose oral reading and reading for comprehension were tested to be on grade level. Subject C drawn in dotted line, Subject D in solid line.

ing for comprehension through operant procedures, but the literature suggests that it is difficult to modify. In contrast, the present data suggest that this behavior can be readily changed. Two comments should be made as to the degree of improvement of Subjects A and B. First, in the parlance of the schools, the performance of on-grade-level subjects was in the "A" to "B" range (85 to 100% correct answers). The Baseline I performance of the below-grade-level subjects was in the "C" to "F" range (65 to 85%), but rose into the "A" to "B" range during the two contingency phases. Most importantly, the comparison of the two sets of subjects shows that the change in the behavior of the below-grade-level subjects brought their reading for comprehension up to a level equivalent to the subjects whose reading for comprehension was on grade level.

It is difficult to relate the present results to

previous studies due to the absence of individual analyses of changes in reading for comprehension in previous studies. If, however, changes were not produced, as the data suggest, two possible explanations for the differences in results should be considered. In the studies by Staats and Butterfield (1965) and Staats *et al.* (1967), the subjects were given consequences for correct answers about material that was at the same instructional or difficulty level as the material on which they were simultaneously receiving oral reading training. Thus, if the subjects were appropriately placed at their oral instructional level, the subjects were asked questions about material they had difficulty reading aloud. In contrast, the subjects in the present study could orally read the material without difficulty over which they were asked comprehension questions. If the subjects in the previous studies had been trained to a given grade level

on oral reading and then trained to answer questions about that level material, different results may have been found.

Another possible explanation for the differences in results of the previous and the present studies is the differences in the reading materials and questions. The passages used in previous studies were longer and the questions were more inferential (e.g., "What do you think Johnny will tell his mother when he gets home?") in the previous studies than in the present one.

The primary question left unanswered by this study is the question of the generalizability of the changes in behavior. Part of the answer to this question is inherent in the design of the experiment: the behavior of the subjects was consistent within each phase of the experiment, even though each passage was different. The results show generalization among different passages of the same form, in other words. Data must be gathered on the more difficult question of generalization to classroom reading and reading where less immediate consequences are provided, however, before applications of this procedure will be justified.

REFERENCES

- Camp, B. W. and Van Doorninck, W. J. Assessment of "motivated" reading therapy with elementary school children. *Behavior Therapy*, 1971, 2, 214-222.
- 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.
- 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.
- Robinson, H. M., Monroe, M., Artley, A. A., Huck, C. S., and Jenkins, W. A. *The new basic readers*. Glenview, Illinois: Scott, Foresman, 1965.
- Spache, G. D. *Diagnostic reading scales*. Montecrey: CTB/McGraw-Hill, 1963.
- 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, 36, 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, 7, 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., Minke, K. A., Goodwin, W., and Landeen, J. Cognitive behavior modification: "Motivated learning" reading treatment with sub-professional therapy-technicians. *Behaviour Research and Therapy*, 1967, 5, 283-299.
- 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.
- Whitlock, C. and Bushell, D. Some effects of "back-up" reinforcers on reading behavior. *Journal of Experimental Child Psychology*, 1967, 5, 50-57.

Received 24 July 1972.

(Revision requested 17 November 1972.)

(Revision requested 19 January 1973.)

(Final acceptance 7 May 1973.)