Published in final edited form as:

Am J Community Psychol. 1979 December 1; 7(6): 689–699. doi:10.1007/BF00891971.

Paraprofessional Reading Tutors: Assessment of the *Edmark Reading Program* and Flexible Teaching¹

Barbara L. Andersen², Barbara G. Licht, Rina K. Ullmann, Sylvia T. Buck, and William H. Redd³

University of Illinois at Champaign–Urbana

Abstract

The purpose of the study was to assess the effectiveness of two strategies for using undergraduate paraprofessionals as tutors for first-grade children with reading difficulties. The two groups of undergraduates taught first graders the same 150 words, and both reinforced (socially and materially) correct performances. However, one group employed a programmed method of word presentation (i.e., the *Edmark Reading Program*), while the other group of tutors was allowed flexibility in the pacing and methods for presenting words. The performance of each reading group was compared to that of a control group who engaged in educational activities not directly related to reading. Thirty-six poor readers were randomly assigned to one of the three groups. Subjects were tutored after school by university undergraduates for an average of twenty-three 40-minute sessions. Children in both reading groups showed significantly more improvement than controls on an oral reading posttest of the 150 criterion words. Treatment effects did not emerge on a test composed mainly of noncriterion words.

Educators have estimated that approximately 30% of American children have difficulty in adapting to school demands (Glidewell & Swallow, 1969). Of particular importance is the large number of children experiencing academic failure; in fact, estimates indicate that one of four students has significant deficits in such a basic area as reading (Allen, 1969). Unfortunately, these problems become more acute as the children move through the educational system. Although it is difficult to determine the exact causal relationship between academic deficits in the early school grades and later mental health problems, the evidence suggests that they are related (Cowen, Pederson, Babigian, Izzo, & Trost, 1973; Zax, Cowen, Rappaport, Beach, & Laird, 1968). For example, one major mental health problem, juvenile delinquency, is frequently associated with reading deficiencies beginning early in the individual's school career (Bachman, 1970; Voss & Elliott, 1974).

Undoubtedly many complex factors underlie academic failure. However, many investigators contend that part of the difficulty stems from the unavailability of teachers and classroom aids to provide the necessary one-to-one attention (Goodman, 1965; Ryan, 1971). Unfortunately, many school districts are unable to afford additional personnel. As a partial remedy, several

¹This research was supported by grant NIE-6-74-0025 from the National Institute of Education awarded to William H. Redd. Appreciation is expressed to Champaign and Urbana School Systems, the Champaign County Head Start Office, Maureen O'Leary, Vernon Savage, Tara Jaffe, Judith Zaruches, tutors, and testers for **their assistance** in conducting this study, to Martin Siegel for his consultation during the early planning stages, and to Kennedy Hill, Robert Linn, and Lloyd Humphreys for their statistical advice. **The authors also wish** to thank Nancy Squires for her **helpful comments** on the manuscript and **Edmark Associates**, Sidney Bijou, and Heed School for providing Edmark Materials.

© 1979 Plenum Publishing Corporation

³All correspondence should be sent to Dr. W. H. Redd, Psychology Department, University of Illinois, Champaign, Illinois 61820.

²Order of first and second authorship was determined by the flip of a coin.

investigators (Cowen, Dorr, Trost, & Izzo, 1972; Hill & Tolman, 1970; Schoeller & Pearson, 1970; Tolor, 1968) have suggested using paraprofessionals to provide individual attention and tutoring. Research by Alden, Rappaport, and Seidman (1975) suggests that paraprofessionals can effect academic gains in children as long as the treatment focuses directly on the child's academic deficits.

If paraprofessionals are to serve as academic tutors, it is important to identify the kinds of teaching materials that are best suited for their use. Programmed materials have been suggested (Ellson, Barber, Engle, & Kampwerth, 1965) because of the relative ease of their administration and the minimal decision-making required on the part of the tutor. In addition, the low error rates typically produced by the sequencing of programmed materials would seem to be desirable in remedial teaching (Bereiter, 1972; Bereiter & Engleman, 1966).

Edmark (1972) is one such program designed for beginning remedial readers (Bijou, Birnbrauer, Kidder, & Tague, 1966). Edmark is particularly appealing for several reasons. First, the amount of training required to administer the program is minimal, as is the amount of prior sophistication required of the tutor. Second, since all lessons are completely planned within the program, tutors need not prepare prior to each session. This enables the paraprofessional to spend the majority of his/her time in direct contact with the target population. And third, lessons are programmed so that new words are introduced in relatively "errorless" fashion, thus making the program applicable for even the slowest, most easily discouraged students.

While Edmark appears promising, it has not been systematically evaluated for use by paraprofessionals nor for normal public school children. (Edmark was initially developed for and tested with retarded children.) While some programmed materials have been successfully used by paraprofessionals in remedial reading programs (Ellson, Harris, & Barber, 1968; McCleary, 1971; Nelson & Peoples, 1975; Staats, Minke, & Butts, 1970; Staats, Minke, Goodwin, & Landeen, 1967), the components vary so much across programs as to prohibit generalization of findings to new programs not yet evaluated. Moreover, while an errorless introduction of new words was seen as a potential advantage, it raised some concerns as well—particularly with a nonretarded population. Specifically, since correct responding requires only minimal effort and attention to the words, it is possible that one could get through the program with little or no retention at the end. Thus, an important question is whether the Edmark method of programming can be effectively used by paraprofessionals with first-grade problem readers.

As mentioned, some programmed materials have been used effectively by paraprofessionals for remedial reading programs. However, what is not clear is how much of the improvements were due to the materials themselves as opposed to the reinforcement (social and/or material) provided by the tutors and the clear specification of criterion skills. Thus, another relevant question was whether or not significant improvements can be effected by paraprofessional tutors who were taught to use reinforcement principles and provided with clearly specified goals, but who were given neither programmed materials nor any structured procedures for presenting the words to be learned (i.e., the material presentation was "flexible").

An important methodological issue in program evaluation studies is the choice of appropriate dependent measures. A common practice is to use standardized reading tests. However, the extent to which the tests employed measure the specific skills taught by the program(s) being evaluated is frequently not made clear (e.g., Nelson & Peoples, 1975; Stephens, Hartman, & Cooper, 1973). When this is the case, interpretation of results becomes difficult. For example, a finding of treatment effects could indicate that either specific or generalized skills had been acquired, depending on how closely the test measured the same skills taught by the program. It is, therefore, desirable to employ two kinds of measures: one which assess the specific skills

taught by the program(s) being evaluated and one which assesses generalization by tapping related, but not identical skills.

Another methodological issue in program evaluation is the choice of appropriate control groups. The typical control comparison in this area has been the no-treatment control (e.g., McCleary, 1971 ; Staats et al., 1970). While scores for this group can be used to determine base-rate gains, they cannot be used to partial out the extent to which improvement of experimental subjects is due to the specific treatment as opposed to the introduction of any type of remedial help. Moreover, since both teachers and children know who is receiving extra training, there is the possibility of expectancy confounds (Rappaport & Rappaport, 1975; Seaver, 1973). Thus, in order to control for attention and expectancy, it is recommended that one include control subjects who engage in extraclassroom educational activities which are not directed at teaching the criterion or criterion-related skills.

In order to deal with the experimental questions and methodological issues raised above, the present study examined the reading improvement of three groups of poor readers: an Edmark reading group, a “flexible” reading group, and an “attention-expectancy” control. Paraprofessional tutors in both reading groups taught the same words, and all tutors administered praise and tangible rewards for correct responses. The three groups were evaluated on two dependent measures: a test of the words taught and a test composed primarily of noncriterion words.

METHOD

Subjects

Thirty-six children (19 males and 17 females; 16 Black and 20 White) ranging in age from 6 years 3 months to 8 years 1 month (mean = 6 years 9½ months) participated as subjects. All were second semester first graders enrolled in approximately 15 classrooms in 7 public schools in Champaign-Urbana, Illinois. All children were referred by their classroom teachers for special help in reading; their pretest scores on the vocabulary section of the Gates-MacGinitie Reading Test-Primary A, Form 1 (Gates & MacGinitie, 1965) averaged in the 25th percentile. All children were from families falling in the lower middle socioeconomic range.

Tutors

Thirty-six University of Illinois undergraduates (12 males and 24 females) received course credit for serving as tutors. They were selected from a pool of approximately 60 applicants on the basis of their interest in working with children and their willingness to apply reinforcement principles with young children.

Tutors were randomly assigned to one of three groups with the constraints that equal numbers of males and females be assigned to each group and tutor-child sex combinations be equated across groups. All groups had at least two pairs of each tutor-child combination. Each tutor was assigned to one child for the entire study.

All tutors were required to read *Parents Are Teachers* (Becker, 1971) and to pass a mastery test on the basic reinforcement principles presented. Each group received approximately 4 hours of pretutoring training, involving discussions, supervisor demonstrations, and practice in presenting tutoring materials. Tutors were also provided written manuals concerning their particular materials.

Throughout the study, two graduate student supervisors assigned to each group spent a combined total of 15 hours observing and giving feedback to individual tutors.

Procedure

Pretest—In order to assess the specific vocabulary taught by the reading programs, the experimenters developed a “criterion-referenced” test (Glaser & Cox, 1968) consisting of the 150 criterion words which were presented in a game-type format. Children were tested individually by White female graduate students. They were asked to read each word aloud and were given 5 seconds in which to respond. No corrective feedback was provided, however verbal encouragement followed approximately every 10th response.

In order to determine whether treatment facilitated acquisition of words not specifically taught, the vocabulary section of the Gates-MacGinitie Reading Test-Primary A (Gates & MacGinitie, 1965) was administered. Of the 40 target words on each of the test’s two forms, only 9 were among the words taught by the reading programs. Form 1 was administered during pretesting in groups consisting of two to seven children. For each group, two White female graduate students were present. One read the instructions, the other passed out the test booklets.

Children were ranked on the basis of the 150 word criterion-referenced test and placed into matched triads (i.e., the top three ... the bottom three). Within each triad, subjects were randomly assigned to one of the three groups. Statistical analyses revealed no significant group differences with regard to sex, race, mean age, or reading achievement vocabulary scores on the Gates-MacGinitie and criterion-referenced tests (one-way ANOVA, all $p > .30$).

Intervention—All children met with their tutors after school for approximately 40 minutes three times each week for 9 weeks. The mean number of sessions was 23 for each group with the exact number of sessions equated within matched triads. When one or two members of a matched triad were absent during the week, make-up sessions were scheduled. If all three members were absent an equal number of times during a particular week, make-ups were not arranged.

Tutors in each of the three groups were instructed to answer all correct responses with social approval (e.g., “good,” “fine”) and to correct all errors. In addition, inexpensive toys and snacks were available to tutors for use as rewards. Tutors and their supervisors individualized the administration of rewards for the children. Some tutors found point systems effective, while others preferred to give a small prize after brief units of work.

Edmark Reading (ER)—Each child in this group was tutored with the *Edmark Reading Program* (1972). Edmark is a commercially available program whose basic material includes 150 word-recognition lessons, each of which introduces one new word. The tutor begins by pronouncing each new word as the child points to and repeats it. The tutor pronounces the word again as it is presented visually along with two “distractors” (i.e., groups of letters or words). The child is instructed to point to the correct word and say it. Finally, the child is required to read the word aloud without prompting when it is presented alone. In subsequent word-recognition lessons the Word is reviewed in phrases and sentences. Each word is also reviewed in supplementary Edmark materials (i.e., matching phrases with pictures, following written directions, and reading aloud stories constructed only of the words taught) which are interspersed with word-recognition lessons in a prescribed order.

Flexible Reading (FR)—Each child in this group was taught the same 150 words as the ER group. However, tutors were allowed complete flexibility in the manner of word presentation and review. The supervisors for this group suggested several methods, such as using flash cards, word wheels, and word bingo games as are described in *Guidelines to Teaching Remedial Reading to the Disadvantaged* (Pope, 1967). During pretutoring training tutors were required to prepare and demonstrate at least one method of their choosing.

Control (C)—A variety of educational activities were used by tutors in this group. Games designed to teach concepts like relative size and relations in time and space were adapted from *Workjobs* (Lorton, 1972). Tutors were each assigned a different project to construct, and each completed project was made available to all tutors.

Posttest—Children were retested approximately 1 to 2 weeks following intervention. For the criterion-referenced test, graduate students unknown to the children and “blind” to the treatment children received served as testers. Each tester was assigned an equal number of children from each group.

Form 2 of the vocabulary section of the Gates-MacGinitie Reading Test -Primary A was administered to groups of three to five children which included children from each tutoring group. For each testing group three adults were present—one supervisor from each of the groups. One supervisor read the instructions while the others passed out and collected booklets. The supervisor reading the instructions was rotated across testing groups.

RESULTS

The three groups (ER, FR, and C) were evaluated on the two tests of reading performance. The .design was a one-way analysis of covariance, with two dependent variables, and one covariate for each dependent variable analysis. Subjects' pretest criterion-referenced and pretest Gates-MacGinitie (Form 1) scores served as the covariates for their respective posttest.

Pretest

Criterion-referenced pretest scores ranged from 0 to 61 of a possible 150. A one-way ANOVA on pretest means (ER = 23.42, FR = 23.25, C = 23.83) revealed no significant group differences, $F(2, 33) = .003, p = .99$. An $F_{\max}$ test, $F_{\max}(3, 11) = 1.14, p > .05$, revealed no significant differences in within group variation ($SD_{ER} = 18.87, SD_{FR} = 17.99, SD_C = 19.23$).

Gates-MacGinitie pretest standard scores ranged from 26 to 53. A one-way ANOVA on pretest means (ER = 44.08, FR = 40.08, C = 40.42) revealed no significant group differences, $F(2, 33) = 1.09, p = .35$. The $F_{\max}$ test, $F_{\max}(3, 11) = 3.26, p > .05$, was also nonsignificant ($SD_{ER} = 6.07, SD_{FR} = 5.43, SD_C = 9.81$).

Analysis of Treatment Effects

The analysis of covariance assumption of homogeneity of regression was not violated, as indicated by nonsignificant F tests for both the criterion-referenced, $F(2, 30) = .49, p = .62$, and Gates-MacGinitie, $F(2, 30) = 1.55, p = .23$, measures.

For the criterion-reference test, the ANCOVA for the adjusted group means (ER = 102.28, FR = 78.50, C = 47.47) revealed an overall significant treatment effect, $F(2, 32) = 14.86, p = .00003$. The above score means reflect average gains of 79, 55, and 24 words for the Edmark reading, flexible reading, and control groups, respectively.

Two planned multiple comparisons were performed to test the experimental questions. To maintain an “experiment-wise” alpha level (Meyers, 1972) of $p < .05$, the alpha level per comparison was set at $p = .025$ ($.05/2$, where 2 equals the number of planned individual comparisons), t -Test comparisons for the Edmark reading against the control, $t(32) = 5.44, p < .01$, and for the flexible reading against the control, $t(32) = 3.08, p < .01$, were both significant.

There was no significant treatment effect for the Gates-MacGinitie test, $F(2, 32) = 1.64, p = .21$. Adjusted standard score means for the three groups were ER = 41.16, FR = 39.07, C = 42.36.

DISCUSSION

Children in both the Edmark and flexible reading groups improved significantly more than controls on the reading test of criterion words. However, treatment effects did not emerge on the standardized test of noncriterion words.

These seemingly inconsistent findings of the two dependent measures, along with similar results obtained by Ellson et al. (1968) and Staats et al. (1967), demonstrate the importance of measuring both criterion and related noncriterion skills. Had the Gates-MacGinitie or other standardized tests of noncriterion skills been the sole dependent measure employed, our conclusions would have been inaccurate. That is, both programs would have been viewed as ineffective. However, both methods were, in fact, effective in teaching the skills they were intended to teach - the 150 criterion words. The lack of generalization to noncriterion words, as shown by the Gates-MacGinitie, was not unexpected since none of the children were taught general word attack skills. Although sight-reading programs have not, as a rule, been found to facilitate acquisition of words not specifically taught (e.g., Ellson et al., 1965; Staats et al., 1967), it has been suggested (Ellson et al., 1965) that instruction of word attack strategies be left until a sight vocabulary has been mastered. Moreover, since a large proportion of English words fail to follow general rules, an effective method for teaching a sight vocabulary is a necessary adjunct to any word attack program.

The effectiveness of both groups of reading tutors adds support to the viability of employing paraprofessionals in remedial education. In particular, college undergraduates provide a readily available pool of easily trainable and motivated personnel. Each semester at the University of Illinois the number of undergraduates wishing to tutor without financial remuneration far exceeds the number of available positions. Given a wealth of this kind of manpower, it is suggested that future research be directed toward finding programs which are best suited for administration by paraprofessionals.

The superiority of the Edmark reading group over the control (i.e., a mean gain of 79 vs. 24 words) provides additional support for the effectiveness of programmed reading "packages," specifically the *Edmark Reading Program*, as well as their appropriateness for administration by paraprofessionals. These results strengthen previous findings (e.g., Ellson et al., 1968; McCleary, 1971; Nelson & Peoples, 1975; Staats et al., 1970, 1967) of the effectiveness of programmed materials in such situations. Although the possible influence of attention and expectancy cannot be ruled out in some of those studies because their control groups received *no* treatment, the present study suggests that these factors were not the most influential in determining their results. However, it is only by comparison with a control such as the one employed here, that one can determine the exact amount of improvement due to the specific treatment employed.

The superiority of the flexible reading group over the control (i.e., a mean gain of 55 vs. 24 words) demonstrates that another viable approach to training paraprofessional reading tutors is to teach them how to apply reinforcement principles, clearly specify the skills to be taught, and provide several alternative methods for presenting the material. Thus, in answer to our second experimental question, significant improvements can be effected by tutors trained in this manner, without providing them with a programmed or structured procedure for word presentation.

Another important question, which goes beyond the focus of the present research, is how the progress effected by "flexible" tutors compares with that of programmed tutors. In the present study, improvements of the Edmark group (i.e., a mean gain of 79 words) exceeded those of the flexible group (i.e., a mean gain of 55 words). The comparison produced a $t = 2.36$, $p = .0245$. These results are only suggestive, however, because the present design does not allow

for a completely controlled comparison. In order to compare the effectiveness of these two methods of word presentation, one needs an assurance that both groups of tutors are delivering equivalent contingencies of reinforcement. To insure this, one must include systematic observation of tutor-child interactions and require that tutors maintain a specified criterion of performance. This is recommended for all future program evaluation studies in the area.

Finally, both research on classroom social behavior (e.g., Ayllon & Roberts, 1974) and sociological theory (Cohen, 1955) would indicate that improvement in reading ability should have important side effects (e.g., less classroom disruption, better social adjustment, improvement in other academic areas). Although without long-term follow-up one cannot determine the broader "mental health" effects of this type of academic intervention, the work of Bachman (1970), Voss and Elliot (1974), and others suggests that the remediation of reading deficiencies might decrease the likelihood of adjustment problems in later years.

In summary, the present study found both groups of paraprofessional tutors to be clearly effective in improving the performance of poor readers. This provides support for each of the programs and for their application by paraprofessionals. Future studies should compare the relative effectiveness of these two methods.

REFERENCES

- Alden L, Rappaport J, Seidman E. College students as interventionists for primary grade children. *American Journal of Community Psychology* 1975;3:261–271.
- Allen J. The right to read- Target for the 70's. *Journal of Reading* 1969;13:95–101.
- Ayllon T, Roberts MD. Eliminating discipline problems by strengthening academic performance. *Journal of Applied Behavior Analysis* 1974;7:71–76. [PubMed: 4465374]
- Bachman, J. Youth in transition. Vol. I. Ann Arbor: Institute for Social Research; 1970.
- Becker, W. Parents are teachers. Champaign, IU: Research Press; 1971.
- Bereiter, C. Conclusions from evaluation studies. In: Stanley, JS., editor. *Preschool programs for the disadvantaged child*. Baltimore: Johns Hopkins; 1972.
- Bereiter, C.; Engleman, W. *Teaching disadvantaged children in preschool*. Englewood Cliffs, N.J: Prentice-Hall; 1966.
- Bijou S, Birnbrauer J, Kidder J, Tague C. Programmed instruction as an approach to teaching of reading, writing, and arithmetic to retarded children. *Psychological Record* 1966;16:505–522.
- Cohen, A. *Delinquent boys*. New York: Free Press; 1955.
- Cowen E, Dorr D, Trost M, Izzo L. Follow-up study of maladapting school children seen by nonprofessionals. *Journal of Consulting and Clinical Psychology* 1972;39:235–239. [PubMed: 5075869]
- Cowen EL, Pederson A, Babigian H, Izzo LD, Trost MA. Long-term follow-up of early detected vulnerable children. *Journal of Consulting and Clinical Psychology* 1973;41:438–446. [PubMed: 4803276]
- Edmark reading program. Seattle: Edmark Associates; 1972.
- Ellson D, Barber L, Engle T, Kampwerth L. Programmed tutoring: A teaching aid and a research tool. *Reading Research Quarterly* 1965;1:77–127.
- Ellson D, Harris P, Barber L. A field test of programmed and directed tutoring. *Reading Research Quarterly* 1968;3:307–367.
- Gates, A.; MacGinitie, W. *Gates-MacGinitie Reading Tests - Primary A, Forms 1 and 2*. New York: Teachers College Press; 1965.
- Glaser, R.; Cox, R. Criterion-referenced testing for the measurement of educational outcomes. In: Weisgerber, RA., editor. *Instructional process and media innovation*. Chicago: Rand McNally; 1968.
- Glidewell, JC.; Swallow, CS. *The prevalence of maladjustment in elementary schools*. Chicago: University of Chicago Press; 1969.
- Goodman K. Barriers to reading comprehension. *Elementary English* 1965;44:853–860.

- Hill C, Tolman R. Tutoring: An inexpensive alternative. *Journal of the Reading Specialist* 1970;10:19–23.
- Lorton, M. *workjobs*. Menlo Park, Calif: Addison-Wesley; 1972.
- McCleary E. Report of results of tutorial reading project. *Reading Teacher* 1971;24:556–559.
- Meyers, J. *Fundamentals of experimental design*. Boston: Allyn & Bacon; 1972.
- Nelson R, Peoples A. A suggested “package” to remediate reading difficulties through paraprofessionals. *Perceptual and Motor Skills* 1975;40:42.
- Pope, L. *Guidelines to teaching remedial reading to the disadvantaged*. Brooklyn, N.Y: Faculty Press; 1967.
- Rappaport M, Rappaport H. The other half of the expectancy equation: Pygmalion. *Journal of Educational Psychology* 1975;67:531–536.
- Ryan, W. *Blaming the victim*. New York: Vantage; 1971.
- Schoeller A, Pearson D. Better reading through volunteer reading tutors. *The Reading Teacher* 1970;23:625–630.
- Seaver W. Effects of naturally induced teacher expectancies. *Journal of Personality and Social Psychology* 1973;28:333–342.
- Staats A, Minke K, Butts P. A token reinforcement remedial reading program administered by black therapy-technicians to problem black children. *Behavior Therapy* 1970;1:331–353.
- Staats A, Minke K, Goodwin W, Landeen J. Cognitive behavior modification: ‘Motivated learning’ reading treatment with subprofessional therapy technicians. *Behavior Research and Therapy* 1967;5:283–299.
- Stephens T, Hartman A, Cooper J. Directive teaching of reading with low-achieving first- and second-year students. *Journal of Special Education* 1973;7:187–196.
- Tolor A. Teachers' evaluations of children in a short-term treatment with subprofessionals. *Journal of Clinical Psychology* 1968;24:377–378.
- Voss, HL.; Elliot, DS. *Delinquency and dropout*. Lexington, Mass: D. C. Heath; 1974.
- Zax M, Cowen EL, Rappaport J, Beach DR, Laird JD. Follow-up study of children identified early as emotionally disturbed. *Journal of Consulting and Clinical Psychology* 1968;32:369–374. [PubMed: 5666136]