DOCUMENT RESUME

ED 064 462 UD 012 816

AUTHOR Rosenshine, Barak; Furst, Norma
TITLE The Effects of Tutoring Upon Pupil Achievement: A Research Review.
PUB DATE 69
NOTE 43p.

EDRS PRICE MF-$0.65 HC-$3.29
DESCRIPTORS *Academic Achievement; Academic Performance; Achievement Gains; Control Groups; Elementary Schools; Experimental Groups; Grade Equivalent Scores; High Schools; Low Achievers; Post Testing; *Program Effectiveness; Secondary Schools; *Test Reviews; *Tutorial Programs; *Tutoring

ABSTRACT

This review focuses on 16 studies in which the effects of tutoring were measured using student achievement, being limited to studies appearing through 1969. Additional studies that the authors subsequently found are considered to still support the conclusions drawn in this review on the general characteristics of successful tutoring programs and successful tutors; and to support the value of directed, structured tutoring when pupil achievement is the criterion. In six of the studies examined in this review, posttest achievement scores for tutored pupils were found to be, in statistical terms, significantly superior to scores of control groups. However, since these results are frequently not presented as grade equivalent scores, it is considered difficult to assess the effectiveness of tutoring in any educationally significant terms. In two other studies that were "successful," nagging questions about the design and outcome measures had been raised; two additional studies which did not have control groups showed one group of middle school pupils as making reasonable progress in reading, and no gains at all in the second project. In five other projects, there were no statistically significant differences between the pupils in the experimental and control groups. Four studies which reported objective data on the effects of tutoring upon tutors were found to be inconclusive; objective measures of affective changes were either nonexistent or showed no significant differences due to tutoring programs. (RJ)

THE EFFECTS OF TUTORING UPON PUPIL ACHIEVEMENT: A RESEARCH REVIEW

1969

Barak Rosenshine
Bureau of Educational Research
Univ. of Illinois
Urbana, IL. 61801

Norma Furst
Dept. of Educational Psychology
Temple Univ.
Philadelphia, PA. 19122

Introduction

Aside from this introduction, this review was completed in 1969, and represents 16 studies which we located in which the effects of tutoring were measured using student achievement. We hoped, and still hope, to complete this review by adding additional studies, but the desire to rewrite this review, other obligations, and the emergance of new studies has prevented the revision.

However, the review was included in the references of the book Children Teach Children (Harper and Row, 1971) (although the review was not included in the text). As a result of the referencing, we have received a number of requests for the paper, and have decided to send a copy to ERIC and make additional copies available. It should be noted that the review is limited to studies appearing through 1969, and the current state of the art might be different.

As far as we can tell, the additional studies we found still support the conclusions we drew in 1969 on the general characteristics of successful tutoring programs (for student achievement) and successful tutors (see pgs. 29-33). Two recent studies (Hamblin and Hamblin, 1972; Neidermayer and Ellis, 1970) continue to support the value of directed, structured tutoring when pupil achievement is the criterion.

At present, our conclusion is that well-structured, cognitively-oriented tutoring programs are relatively few, but when they occur, there are usually measurable achievement benefits to the pupils. The majority of tutoring programs apparently do not have these characteristics but consist of less structured, helping, affective interactions. In these "softer" situations, the anecdotal reports are that the tutors and the pupils develop increased pride, positive attitudes toward self and school, enhanced self-image, and greater patience. Such ends may be sufficient for some readers, but no reader should believe that increased pride is equivalent to increased reading ability until the data are in.

Effects on tutors. Within the current review, we do not have a section on the effects of tutoring upon tutors because we found only two studies which presented hard data. (In most studies, the t ors were college students, or public school students who did not need tutoring.) In the text, Children Teach Children, by A. Gartner, M. Kohler, and F. Reissman, there are anecdotal reports of benefits to tutors but no reports of achievement gain which are not included in our review. Thus, although we would have liked to write more about the effects of tutoring upon the tutors, there is insufficient evidence for such a review, at present.

THE EFFECTS OF TUTORING UPON PUPIL ACHIEVEMENT: A REVIEW OF RESEARCH

Barak Rosenshine

Norma Furst

Temple University[1]

The use of tutors and classroom aides has frequently been advocated as one method for improving the academic performance of low achieving pupils (Cf. Passow, 1967; Goldberg, 1967). By all chat is reasonable, such procedures should measurably help pupils. The tutor can attend to the particular difficulties of his pupil, allow him a good deal of practice, provide corrective feedback, and provide reinforcement in the form of praise and assessment of progress. We might expect also that the positive effects of tutoring would generalize, so that the pupil might grow in measures of aspiration and self-esteem, as well as improve in both attention and participation in his regular classroom activities.

Tutoring programs, particularly those designed for low achieving pupils, have spread widely in recent years. In many schools, parents and college students in the community, as well as older pupils, have spent a few hours each week tutoring one or two children. A review of published results of tutoring programs appears appropriate in light of their seemingly wide acceptance.

---

[1] Robin Nelson, Paula Plourde, and Jean Conwizer helped in the review of research; Robin Nelson and Barbara Rosenshine provided invaluable editorial assistance. We are also grateful to those investigators who responded quickly to our requests for reports and/or additional information.

Tutoring, as almost any educational practice, defies precise definition. Not only are the hours and the teacher-pupil ratio modified to meet changing events, but the content, the materials, and the instruction differ widely even within the same school. At the minimum, a tutoring situation would be one in which no more than three or four school-age pupils (frequently one) are tutored by someone other than the regular teacher for one to four hours a week.

In compiling this review, thirteen projects were uncovered in which school age students were tutored and objective data were collected and reported. Ten of the projects included data on control groups; three investigations provided analyses of pretest and posttest data only for tutored students.

In view of the amount of publicity given to tutoring, thirteen projects reporting data seems to be a rather small number. It certainly seems an underestimate of the number which should have been reported. However additional studies could not be found in educational journals, or in the ERIC collection. Inspection of reviews of research in dissertations and reports on tutoring indicated that other investigators were no more successful in locating additional tutoring studies that analyzed objective data than we were.

However, it should not be surprising to find that the amount of controlled research on tutoring is very small in proportion to the amount of

tutoring that is taking place. The selection of experimental and control groups is a difficult procedure, and many teachers and tutors are reluctant to deprive a pupil who apparently needs tutoring of that additional instruction by placing him in a control group. In addition, the problems of administering pretests and posttests are wearisome, and testing takes up class time. Finally, controlled objective testing appears unnecessary to many in view of the overwhelmingly favorable reports given both by the tutors and the teachers involved in tutoring programs.

Control groups are important in tutoring studies because we would normally expect that any group of pupils, whether tutored or not, would make some progress over a summer or a semester, and that this progress would be reflected in the results of a "correlated t-test." In addition, it is rather hazardous to project "expected gain" on any standardized achievement test because these tests do not have longitudinal norms. Nevertheless, three tutoring studies which did not include control groups are reviewed here because the number of tutoring studies with control groups is so small.

ORGANIZATION OF STUDIES

In this review, the tutoring projects have been classified as "successful" or "unsuccessful." "Successful"includes all studies in which at least one of the tutoring objectives was achieved [illegible] as measured by objective tests. In one project (Ellson et al., 1968) multiple classification conditions with differing results resulted in the study being discussed under two classifications.

Successful Studies with Control Groups

Reading in Grade 1. In a series of studies, Ellson and his associates (1968a, 1968b) have examined the effect of programed tutoring upon the reading achievement of low achieving first grade pupils. Under the programed tutoring condition, the tutors led the children through one sight-reading program and six comprehension programs by following specific steps outlined in each program. The tutors were not teachers, and they had no previous training in this area. The program placed strong direction upon the instructional behaviors of the tutors.

In the first study of programed tutoring (Ellson et al., 1968), selected first grade children were placed randomly into one of three condition groups: (a) two fifteen minute daily sessions of programed tutoring, (b) one fifteen minute daily session of programed tutoring, or (c) no tutoring. The experiment lasted for 28 weeks and took place during school time, but it was given in addition to the regular reading instruction. Testing was conducted in September, January, and June.

On the basis of the June posttests, pupils who received programed tutoring had significantly superior reading scores to those of the control ($p < .01$), but these results were influenced almost exclusively by the group which had two programmed tutoring sessions daily. The scores of pupils who had one programed tutoring session daily were not usually statistically superior to the scores obtained by controls.

During a subsequent year, only one session of programed tutoring was used for the experimental group. Posttest data for that group were compared with those of the controls (Ellson et al., 1968). Although the statistical significance of the results is not given in the short

report, the tutored pupils achieved posttest scores which were superior to those of the controls.

Reading for Fourth and Fifth Grade Pupils. Cloward (1967) attempted to assess the effects of the Mobilization for Youth program in New York City. The final sample for analysis consisted of 356 experimental subjects and 157 control subjects who were in fourth and fifth grade classes and were reading below grade level as measured by the New York Tests of Growth in Reading.

Students were tutored for one or two afternoons a week for a five month period. Tutoring was done by high school students in 11 tutorial centers. A certified teacher supervised each of these centers and provided the tutors with two hours of training each week. Tutoring sessions were described in this way:

> "By the end of the second month, the typical tutoring session consisted of 30 minutes spent on homework, 30 minutes on reading, 15 to 30 minutes on games and recreation, and 15 minutes for refreshments, roll-taking, and other non-tutorial activities."

Differences between the groups in reading growth were analyzed by subtracting pretest raw score from posttest raw score and correcting these raw score differences using analysis of covariance in which pretest scores, sex, and school grade were included among the covariates. The adjusted mean difference scores were only slightly different from the unadjusted mean difference scores.

Taken as a group, the tutored subjects made a gain that was slightly superior to that of the controls, but this result was not statistically significant. A second analysis was made for (a) those

pupils who were assigned to be tutored two afternoons a week, (b) those pupils in the same centers receiving tutoring only one afternoon a week, and (c) the appropriate controls. In this analysis, the adjusted difference scores of pupils who were tutored two afternoons a week were significantly ($p < .05$) superior to those of the controls. There were no significant differences between those tutored one afternoon a week and the controls.

When the results are expressed as grade equivalent scores, those tutored two afternoons a week averaged six months gain in the five month period, those tutored one afternoon a week averaged five months gain, and the controls averaged 3.5 months gain. At this rate of progress, the most successful group would have to continue attending the tutoring centers for at least four more years before they reached grade level achievement in reading. It should be noted that following the evaluation by Cloward, none of the reports on this program contained any data on student growth (Deering, 1968, 1969).

Reading and writing--grades 4, 7, and 10. A most comprehensive series of tutoring experiments has been reported by the Logan-Cache school districts of Utah (Logan City, 1968; Shaver, 1969). Their research into tutoring was aimed at finding what differences, if any, might be evident between tutored and non-tutored students, among different grade levels, and with differing tutor-tutee arrangements. Further, they were interested in results of "delayed testing," or comparisons between tutored and non-tutored students one and two years after completing the program.

The tutors included former teachers, graduate students, and housewives who were prepared in a ten-day workshop (9 to 3) with heavy emphasis on training in specific skills and materials suitable for developing specific skills with underachievers. Tutoring sessions were held for a full school year.

Third, sixth and ninth grade students were given the California Test of Mental Maturity and STEP tests of reading and writing ability to determine their eligibility for tutoring during their fourth, seventh, and tenth grades. With the students' scores on the CTMM as the criterion, overall correlations between CTMM scores and [illegible] on the two STEP tests were used to predict how well each student [illegible]e doing on the STEP tests. This measure of "potential" was used as the criterion. Those children whose scores on the achievement tests were from 5 to 20 points below their predicted scores were randomly assigned to either (a) a one tutor - one student condition, (b) a one tutor - three students condition, or (c) a control group that remained in the regular classroom. Objective evidence of the effectiveness of the tutoring was determined by analyzing the results of STEP tests given in midyear and at the end of the year, the amount and quality of reading and writing done by the pupils, and teachers' grades.

The research pattern was replicated for two years. The third-year experiment dispensed with the control group and substituted a one tutor - five student situation in its place. The researchers were also interested in whether the gains of the tutored students over the control students continued from one to two years after the experience. Therefore,

they readministered the STEP test in the Spring of 1969 to the students who had been involved in the 1966-67 school year program. In addition, students' grades at the end of the 1967-68 school year and at the end of the first semester of the 1968-69 school year were analyzed.

STEP test data were analyzed for the first year, using analysis of covariance. The results indicated clear, statistically significant differences between control and experimental groups, favoring the students who had been tutored. It also appeared that the difference became increasingly greater from the fourth to the seventh to the tenth grades. This same picture of effectiveness was found in comparing the students' scores on a reading comprehension test developed by the tutors at each grade level.

Chi-square analysis of comparisons of the number of tutored students who attained their potential or better during tutoring and the control students who attained their potential or better also clearly favored the tutored groups.

There were no significant differences between the one-to-one groups and the one-to-three groups, and both groups did better than the controls. Third year analysis, which included a condition of one tutor - five students, showed n.s.d. among the three tutoring arrangements. However, there were far too few 1 to 5 teams to allow for anything more than tentative conclusions.

The results of the analyses for the second year were similar to that of the first year, except that the differential effectiveness of tutoring at different grade levels was not as marked in terms of the number of students reaching potential or better.

Analysis of school grades indicated no clear pattern favoring the tutored students in the fourth grade. However, students tutored as

seventh or tenth graders do appear to have a significantly superior mean grade in English, science and social studies--those subjects most closely related to the tutoring experience.

The results of the analysis of the "delayed" STEP test data indicate that the magnitude of the F ratios decreased, with statistically significant differences favoring the tutored groups remaining for the seventh and tenth grade levels. However, the differences between the tutored and control group means for the students tutored as fourth graders were no longer significant.

Shaver and his associates should be commended for the thoroughness of their approach and the attempt at answering multiple questions about tutoring and its effects. However, we are concerned about the use of a measure defined as "pupil potential," and the fact that grade equivalent scores were not presented for any of the groups. A criterion labeled "potential" is not identical to grade equivalent scores; it apparently means that if a student has an IQ below 100 and is reading below grade level, this is an acceptable situation as long as he is reading "above potential." A situation in which a student with a measured IQ of 80 is reading above "potential" but below grade level may not be acceptable to many readers.

Low Achieving Fifth and Sixth Grade Pupils in Arithmetic and Word Knowledge. Glatter (1967) studied the effect of nine weekly two-hour tutoring sessions upon the arithmetic and word knowledge scores of 60 underachieving fifth and sixth grade pupils. A second, untutored group served as controls. The tutors were 60 college juniors or seniors.

The low achieving pupils had an average IQ of 89 (California Test of Mental Maturity) and were at least one year below national norms on the Iowa Tests of Basic Skills.

> "Each session consisted of formal, individualized instruction in the basic 'three R's,' focusing attention upon the pupils' particular weaknesses. This hour was followed by one hour of singing, group games, refreshments, and art work" (Glatter, 1967, p. 19).

The particular tutoring procedures used were not specified, but Glatter adds that "standard reading and arithmetic texts were used regularly . . . tutors often distributed mimeographed sheets listing words, multiplication tables, and problems in arithmetic" (Glatter, 1967, p. 23).

The effects of the program were studied using the arithmetic computation and word knowledge subtests of one version of the Metropolitan Achievement Tests--Intermediary Level (MAT). This test was given during

the first and tenth weeks of the program. Because there were no significant differences on the pretest, the major hypotheses were tested by comparing the posttest scores for the two groups. However, only 21 of the original control pupils took the posttest, and only those experimental pupils who attended seven or more of the nine tutoring sessions and took the posttests were included in the analysis. This restriction reduced the size of the experiemntal group tc 40 of the original 60.

Based upon the raw posttest scores, the experimental group was superior to the control group on the arithmetic computation subtest, but not on the word knowledge test. Although grade-equivalent scores were not used in the analysis, Glatter states that the experimental group progressed from "an average of 4.2 to 4.7 grade level on the arithmetic computation and from an average of 4.5 to 4.7 grade level on tested word knowledge" (Glatter, 1967, p. 28).

The grade equivalent scores indicate that the program was a qualified success, but we must note that the tutored fifth and sixth grade children were still substantially below grade level at the end of the program. In word knowledge, the tutored pupils (who represent only two-thirds of the original sample) made two months progress in approximately three months, a rate which is similar to their previous record. The gain in arithmetic computation is almost double the progress which might be "expected." However, the arithmetic computation subtest of the MAT requires a pupil only to compute a series of arithmetic operations; there are no words or word problems in this test. Of all the tests in any standard achievement test, arithmetic computaion is most similar to the tutoring situation. Tests of word knowledge, reading, or arithmetic

problem solving require more general knowledge, whereas a test of arithmetic computation is much more specific, or "factual." Therefore, the gains in arithmetic computation as a result of tutoring are not surprising; nor is the lack of gain in word knowledge scores. It is unfortunate that the reading subtest of the MAT was not also used in the study.

Tutoring of Low Achieving Freshmen in High School English by Low Achieving Seniors. Werth (1968) studied the effects of tutoring in English on both tutors and tutees, using 32 high school freshmen classified as "low achievers" and 30 "low achieving" high school seniors who served as tutors. Tutoring was conducted during the regular English class period, one day a week, for one school semester. Thirty-two seniors and 32 freshmen in the same English classes served as controls. Criterion measures were difference scores (posttest minus pretest) on the Gates Diagnostic Reading Tests and the Language section of the California Survey of Academic Achievement Test.

During the tutoring sessions the freshman read a short story and was helped by his tutor to complete a study guide on the material. The study guides were prepared by the investigator. "The study guides for literature included vocabulary, fill-in comprehension checks, short answer comprehension and inference questions, and short (three or four sentence paragraphs) essay questions" (Werth, 1968, pp. 39-40). More specific information on the content or procedures of the tutoring sessions was not given in this report.

The effects of the tutoring program were analyzed using pretest minus posttest difference scores (raw scores) for the experimental and control freshman students. Those freshmen who were tutored made slightly better gains on the reading, language, and spelling tests, but the differences were statistically significant (p .05) only on the reading tests. The report did not express pretest or posttest scores in grade-equivalent forms.

Mathematics and Foreign Language in High School. Lundberg (1968) brought high achieving high school students and students who were experiencing difficulties in mathematics and foreign languages together in a tutoring relationship for a period of one semester. Six Southern California high schools were involved in the project. In each school a supervisor of tutoring was appointed by the principal. Five of the six supervisors were counselors. The leader of the program was responsible for organizing an orientation conference with each tutor. These sessions were short and revolved around the theme, "A good tutor is a careful listener, and asks many whys. The tutor should work with his student and not preach at him" (p. 100). The supervisor periodically visited the tutoring pairs to give encouragement.

Seventy-four underachieving high school students volunteered to be tutored. They had to personally seek help. The student who was to be tutored also determined the frequency of the tutoring sessions, the length of each session, and the point at which tutoring was terminated.

A comparison group consisting of 101 students who were considered by their teachers to be "like the tutored group" was employed for

statistical analysis. None of the comparison students had volunteered for tutoring, nor had any shown any indication of the self-initiative which was a vital aspect of the tutored group. This difference must be kept in mind when discussing the results of the study which indicated that the percentage of tutored students with semester grades of C or better was significantly greater than for the comparison group. No mention is made of whether or not the teachers who gave the grades knew to which group the students belonged. Lundberg also found that students who had elected to have after-school tutoring, and thus had to pay for the experience, showed significantly greater grade improvement than students who did not pay.

Successful Studies Without Control Groups

Reading in Grades 4,5,and 6. Hassinger and Via (1969) report the results of a tutoring study done in six school districts in Los Angeles County. The tutors were "disadvantaged" high school students who were two to three years retarded in reading, in addition to school dropouts and unemployed high-school graduates. They tutored fourth, fifth, and sixth grade underachieving elementary school students in reading in two-hour blocks for six weeks.

A pre-service training period was held in which reading specialists introduced both the teacher-supervisors and the tutors to basic reading materials. Each tutor was given instruction in the use of audiovisual equipment and in the practice of word games and other "high interest devices." The teacher-supervisors also spent four hours per day for four days planning with the tutors and physically organizing each classroom for the tutoring experience.

Hassinger and Via report a mean growth for all tutees of 4.6 months in reading during the six-week program period on the Stanford Reading Test.

Pre-College Tutoring. Silver (1967) conducted an intensive six-week summer program in reading, writing, mathematics, and language arts for 27 high school graduates who were admitted to Bakersfield College (a two-year college in California) but who scored below the 11th centile on the SCAT and the English Classification Test. The group met for four class periods a day, five days a week.

The results in reading were evaluated using the California Achievement Tests. Both the pretests and the posttests were utilized in the analysis. On the reading test, the mean score was raised from 8.0 at the start of the summer to 8.4 at the end of the summer. At that rate, it would take about a year under the same circumstances to attain grade level.

Unsuccessful Tutoring Studies with Control Groups

IQ Change in Kindergarten. In a complex study which had as its main focus change in teenage tutors, South-Western City School District (undated) reported results of tutoring on the tutees' performance on an intelligence test. Seventh graders had been given a special course in child development and were employed under highly supervised conditions to work in a one-to-one or one-to-two relationship 40 kindergarten children during the younger children's regular class time.

It had been hypothesized that one effect of the program would be a change in the kinde ;artener's pre-reading communication skills. To test this hypothesis, the Peabody Picture Vocabulary Test was given to the experimental group of children and to a control group of kindergarten children from another school as a pretest and a posttest. At the beginning of the program, both groups of children had comparable mental age scores (5.5 and 5.6) with a chronological age of 5.6. The posttesting revealed a change in IQ of 2 points for the control group and less than one point for the experimental group.

The researchers try to explain their findings by suggesting that the Peabody Picture Vocabulary test "is not a very reliable test of intelligence, or simply does not reflect the change in knowledge one might expect from this type of experimental program." To further confound the interpretation of their results, they report that both groups had been exposed to teacher aides during the experimental program, and that differences might therefore have been masked due to that unexpected variable. Observers' and teachers' perceptual, anecdotal reports on the kindergarten children in the experimental group did indicate that they "saw" the tutored group as having made progress during the experiment.

First Grade Reading. In the first study by Ellson and his associates (1968) two of the experimental conditions involved regular, or non-programed tutoring, in which the first grade pupils were tutored for 15 minutes (a) once a day, or (b) twice a day. Although children who received the regular, or "directed" tutoring had posttest scores

which were superior to those of the control group, none of the differences--in either the one-session or two-session condition--was statistically significant.

Tutors in the"directed" tutoring condition received the same number of hours of training as the tutors in the more successful programed tutoring condition reported above: four three-hour sessions before tutoring, and three three-hour training sessions during the school year but were not required to use programed behavior or the special programed materials. Ellson commented that he was surprised at this result, because the tutors in the "regular" program received extensive training directed towards the development of the reading skills of first graders. Specific training was given in developing reading readiness, skills of visual and auditory discrimination, left-to-right sequence, rhyming words, and visual motor skills. Yet, students in this regular tutoring program did little better than control students who did not receive the tutoring.

Tutoring of Second and Third Year Pupils. Kirk (1966) evaluated a two-year program of tutoring in which 44 children were tutored the first year, and 27 children were tutored the second year. Those children selected for tutoring, and the controls, had verbal ability scores ranging from 80 to 100, and had Stanford Reading Test pretest scores from 1.1 to 1.9. Tutored children were divided into three groups: (a) those who received more than 20 hours of tutoring (during school time) throughout the semester, (b) those receiving between 10 and 20 hours of tutoring, and (c) those receiving less than 10 hours of tutoring. At the end of the first year, the non-tutored pupils had significantly higher posttest scores ($p < .001$) after the scores were adjusted for pretest scores. At the end of the second year, there were no significant differences.

In neither year was there any meaningful correlation between minutes of tutoring and gain in posttest scores (r's = .10). The tutoring procedures were not clearly specified in the report.

Middle School Reading and Arithmetic. One requirement of an educational psychology course for college juniors was that they spend a minimum of one hour a week tutoring a student in a public school. The amount of tutoring instruction which each Junior received varied with his college instructor; however, all college students tutored at least two hours each week. Rosenshine and Furst (1969) attempted to evaluate the effectiveness of this program by comparing the scores of middle grade students from public schools who received tutoring with the scores of similar control students. Scores on city-wide Iowa Tests of Basic Skills were used as pretest and posttest scores. There were no significant differences between the 18 tutored and 18 nontutored pupils on the pretests, and nonsignificant differences persisted on the posttests.

A replication employing better procedures of randomization and utilizing students from only one elementary school situation yielded similar null results (Furst, Rosenshine, and Mattleman, 1970).

High School Achievement. Weitzman (1965) reported a study in which one teacher selected certain pupils for tutoring by high school students and compared the changes in these pupils with those of similar pupils in the same classroom. Interestingly, the teacher reported significant changes in their classroom behavior of the tutored pupils. However, there were no significant differences between tutored and control pupils on the teacher's own tests.

Summer Tutorial Program in Grades 3, 4, and 5. Grannick (1968) evaluated the results of two summer tutorial programs conducted in 1967. One program took place in Philadelphia; the other, in Newark, New Jersey. In both programs, the tutors were primarily students, 14 and 15 years of age, who were reading below grade level.

Those who supervised the tutors (certified teachers and teacher-aides) received one week of training, and the tutors received a week and a half of training. The training procedures differed in the two cities. In Philadelphia, the training for supervisors was focused jointly on affective, or sensitivity training, and on help in remedial reading provided by a specialist from the Board of Education. The trained teachers then provided the training for the tutors.

In Newark, the supervisors (all of whom were untrained women from the community) received training in a "highly structured and technical method for training indigenous mothers to tutor children who had reading difficulties" (p. 29), and the tutors received five full days of similar training. Once the program began, the tutors in both cities met for approximately 10 hours a week for formal and informal training.

No evaluation was made of reading improvement for the pupils tutored in Newark. In Philadelphia, the reading and word knowledge subtests of the MAT were administered as pretests and posttests. Attrition rates were very high, and complete data were available for only 51 of the 588 pupils who enrolled (pp. 24 and 78). These 51 pupils represented a pooling across the six schools involved in tutoring. Only raw scores are presented in the tables. There was no significant change from pretest to posttest for either of the subtests. The pretest and posttest means were almost unchanged.

Summary

In summary, six studies were presented in which posttest achievement scores for tutored pupils were found to be, in statistical terms, significantly superior to scores of control groups. From these studies one might make some hesitant claims for the value of tutoring over no additional experience. However, the results are frequently not presented as grade equivalent scores, so that it is difficult to assess the effectiveness of tutoring in any educationally significant terms. Whenever grade equivalent scores were available, short-term tutoring did not seem to produce either statistically significant gains or the kind of advances toward attainment of grade level that might be hoped for.

In two studies that were "successful", nagging questions about the design and the outcome measures have been raised. Two additional studies which did not have control groups showed that one group of middle school pupils made "reasonable progress" in reading, and no gains at all were made in the second project. In five other projects, there were no statistically significant differences between the pupils in the experimental and control groups.

Changes in the Affective Domain

The preceding review concentrated only on changes in achievement measures. However, no review would be complete without a discussion of changes in the affective domain related to tutoring experiences.

In several studies, tests of pupil attitudes toward education, reading, or perception of self were given. None produced significant results.

Cloward (1967) administered a questionnaire which included items concerning the pupil's educational and vocational aspirations, and his attitudes towards school. According to his analysis, the tutorial program had no measurable effect upon pupils' attitudes and aspirations. Nor did pupils receiving the most tutoring, or pupils making the highest gains in reading, have significantly higher posttest scores on the attitude questionnaire; nor were any of the attitudes or aspirations measured predictive of reading improvement.

Glatter (1967) administered a School Attitude Questionnaire to the 38 pupils in the tutored group. The questions focused upon the pupil's liking for various school subjects, the school, and the values he and his family generally placed upon education. The means for pretest and post-test were almost identical. Glatter completed an item analysis of the questionnaire and concluded that a negative trend in attitude towards school could be discerned. He concluded that this deterioration may represent a more realistic appraisal by the student of his own knowledge and standing as a result of his tutoring experience.

Glatter also found that none of the initial scores on (a) attitude towards school, (b) self-concept, or (c) pupil's own social desirability were positively or significantly related to pupil achievement within the program.

Grannick (1968) used a third survey instrument on reading attitudes, and analyzed the results for Newark and Philadelphia separately and found no significant differences (or even trends) between pretest and posttest. Comparison of pretest and posttest scores on all 25 items likewise indicated no significant changes (Grannick, 1968, p.77).

Rosenshine and Furst (1969) administered the Brookover Self-Concept Inventory as a posttest to both the tutored and nontutored middle school children in their study. No significant differences were found between the groups.

Teacher and observer perceptions in the form of checklists or anecdotal reports (Weitzman, 1965; South-Western City School District, 1969; Kirk, 1966; Rosenshine and Furst, 1969; etc.) all have been used to support the argument for benefits to the tutees in the social-emotional realm. However,there has been a dearth of results when attempts have been made to measure these effects with more precise instruments.

The overall results of the objective affective measures are far from encouraging. Even though different test instruments were used in idfferent projects, data analysis showed no significant differences from pretest to posttest, with no exceptions.

The lack of significant differences becomes even more striking when the identical reports cite subjective evidence from tutors and from teachers indicating atrong positive changes in the attitudes of those

being tutored. We must conclude this section by noting the strange and irreconcilable difference between the objective measures of pupil attitudes (including a variety of tests) and the subjective reports from those engage in the tutoring programs.

The Tutors

Some of the investigators who have studied tutoring have also studied (a) the effect of tutoring upon school-age tutors, and (b) the characteristics of an effective tutor. There are fewer studies on tutors than there are on the effects of tutoring on the tutee, and therefore any conclusions are severely limited by the inadequate number of investigations.

The Effect of Tutoring upon School-Age Tutors. Although there has been a good deal of testimony in favor of the effects of tutoring upon school-age tutors ("cross-age tutoring"), there has been little objective research in this area, and the few results are difficult to interpret.

From the eligible 10th and 11th grade students who applied for positions as tutors, Cloward (1967) randomly selected 155 as tutors and told 72 others that they would be offered tutoring positions the following year. Eligible tutors were (a) 16 years old or older, (b) not in danger of failing their school work, and (c) no more than three years below grade level in reading. Of the original 155 tutors, 37 did not complete the program, leaving a final sample of 97. Twenty per cent of the control subjects were lost, leaving a final sample of 57. On the pretest, the

experimental and control groups were quite comparable, reading an average of 7 months below grade level.

The data were analyzed by subtracting pretest from posttest scores, and these differences were adjusted using analysis of covariance with pretest reading level and the Quick Word Test among the covariates. There were significant differences favoring the tutors on reading comprehension, directed reading, and the total test score.

Expressing the scores as grade equivalencies, in the seven months between the pretest and the posttest, the control group showed a mean growth of 1.7 years; the experimental group gained 3.4 years. Increments of this size are difficult to interpret. Cloward claims that "a substantial portion of the increase for both groups was due to their increased familiarity with the complex directions for taking (the alternate form of) the test" (Cloward, 1967, p.22). Using this interpretation of the results for the control group, one might assume that high school pupils reading an average of six months below grade level can be brought to grade level and beyond merely by giving them a second form of the test.

Despite the difficulty in interpreting the results as grade equivalent scores, those pupils who served as tutors improved significantly more in their reading ability than those who served as controls. However, the generality of this finding is restricted by the high attrition rate.

In the study by Werth (1968), the low achieving senior student tutored the freshmen one period a week fro an entire semester. Both the tutors and the control students spent one period each week studying the

material which the freshmen were to learn the next day during the tutoring session. For the seniors who served as tutors, there were no significant differences between experimental and control groups on the reading, language, and spelling tests. However, on the language tests the differences in favor of the tutors were significant at the .10 level. Unfortunately, raw scores were used in the analysis, so that we cannot make any estimate of gain expressed in grade-equivalent scores.

In the study by Grannick (1968), the data on the tutors in Newark were analyzed separately from those on the tutors from Philadelphia. (It should be recalled that there was no control group.) The Iowa Silent Reading Tests (the same tests used by Cloward) were administered. Different forms were administered as pretests and posttests.

In Philadelphia, the tutors--who were 0.4 years below age level on the pretest, gained one year during the seven-week program. On a correlated t-test, these gains were not significant. In Newark, the tutors began at a lower level, reading 2.9 years below age level. They gained 3.4 years during the seven-week period, and the differences were statistically significant.

Grannick also reports that the Newark program was completely run by inner-city parents, and that there was community motivation to achieve success. "There was some indication that the tutors were concerned that poor performance on their parts might result in loss of the program . . . .In the post-testing sessions it was observed that they attempted many more items than in the pretest" (Grannick, 1968, p. 75).

Hassinger and Via (1969) report that their disadvantaged high school and post-high school tutors had a significant mean gain of eight months on the Nelson Denny Reading Test after the six-week tutoring experience. They concluded that, although the teacher-supervisors ranged from reading teachers to physical education instructors and the districts varied both socioeconomically and geographically, the effect of the program on all groups of tutors was positive.

Hassinger and Via (1969) report an interesting "measure" of change within their group of tutors who had been selected from a population of both socioeconomic and scholastic disadvantage. Although no data were collected, the investigators noticed an evidence of change in attitudes discernable in a change in physical appearance after the second week of the program. Tutors who had worn beards, hair in curlers, and extremely informal dress began to wear more conventional clothing--such as hose for many of the girls and white shirts and neckties for a group of five male tutors.

The South-Western City School District of Ohio (undated) reports an interesting and complex study evaluating the demonstration phase of a teen tutorial program. In their project, seventh grade students worked as tutors for kindergarten pupils. Parents and community workers were also included in the program. The report includes objective data on both the kindergarten students and the junior high school tutors.

Forty seventh grade students who met the standards used by the Office of Economic Opportunity for low income groups, had an IQ of 80 or above on the California Test of Mental Maturity (administered during the sixth grade), and were known to be free of any known severe handicap

made up the experimental group of tutors. Control students were chosen from another junior high school and matched to the experimental subjects on six criteria.

The experimental group of tutors was given a specially designed course in child development with special emphasis on social relationships. They also used the kindergarten situation for tutoring and as a field laboratory for their course work. Both the experimental and control groups were given a series of objective tests as pretest and posttest. These included a special group of tests written by the research team to assess cognitive knowledge of selected areas of child development and social relationship principles, the California Test of Personality, and the Michigan State Self-Concept Test (Brookover, 1962, 1965).

There were statistically significant differences between the experimental and control groups on only one test--the research team's objective subtest for knowledge of five-year development. All other objective tests of knowledge and affective measures showed no significant differences between the "teen" groups.

Subjective measures by junior high school teachers, the kindergarten teachers, parents and the teens themselves indicate, for the investigating team, that the tutoring students did benefit from the experience. These data were mainly "anecdotal" in nature, with no "base-line" data taken, and with no control group measures with which to compare them.

Lundberg (1968) reports percentages taken from questionnaires given his tutors. They seem to report that the tutors perceived that their experience gave them some improvement in their knowledge of the subject and their ability to work with pupils. The majority of the students agreed that they would like to tutor again.

Summary

In summary, the four studies which report objective data on the effects of tutoring upon school-age tutors are inconclusive. In studies which used control groups, there were significant gains for the tutors in one (Cloward, 1967) and not in another (Werth, 1968). In the two studies reported by Grannick which did not use control groups, there were significant differences in only one case. The Southwestern City School District of Ohio (1969) study showed significant differences in only one of a host of objective measures given the experimental and control groups. This measure probably related most strongly with the subject matter introduced to the experimental group in its special class experience; with the more general knowledge and affective measures showing no significant differences.

Objective measures of affective changes either are nonexistent or show no significant differences due to tutoring programs. However, subjective and anecdotal data are used widely to support the efficacy of tutoring for the tutors.

There are three issues that complicate the interpretation of the objective results. First, the positive results in Newark may be confounded by the fact that it was an exceptionally highly structured program, and by the strong community pressure on the tutors to succeed. Either or both of these variables may have been influential in the results. Strong community pressure appears to be an important variable, but the existence of such pressure restricts the generalizability of the Newark results. These results were also not replicated in the Ohio study, in which parents and community resource people were also heavily involved. Second, the

age-equivalent scores on the Iowa Test of Silent Reading are difficult to interpret. The control group in the study by Cloward gained 1.7 years in seven months; the non-significant gains in the study by Grannick were 1 year and seven weeks. The Hassinger and Via study reported a gain of eight months in six weeks, which was significant in terms of pretest and posttest, but there was no control group.

Third, the study by Werth suggests that practice without tutoring may be as effective as tutoring itself. In this study, the control group which studied the tutoring materials made as large a gain as the tutors who both studied the materials and tutored. Such results, if replicated, might suggest modifications in the traditional instructional program for low-achieving pupils.

Characteristics of Successful Tutoring Programs

Hawkridge and his associates (1968) prepared a review in which 18 well designed, successful programs for producing cognitive gains in disadvantaged children were compared with 27 matched unsuccessful programs. After completing the review, their major recommendations for establishing sound programs might be summarized as follows:

1. Careful planning, including clear statement of academic objectives
2. High intensity of treatment with instruction and materials closely relevant to the objectives
3. Individual attention to pupils' learning problems

Whereas the unsuccessful programs generally contained these elements, the unsuccessful programs were more diffuse in their objectives, attempting to provide a variety of enrichment services. In the unsuccessful programs, more time was spent on cultural activities, and less time on academic activities.

These characteristics apply, in general, to the successful and unsuccessful programs described above. The programed tutoring packages developed by Ellson and his associates were highly structured; specific tutoring materials were prepared and studies beforehand in the program developed by Werth (1968), and the significant results in this study were in reading comprehension, the area most directly related to the instructional materials. By comparison, their was less structure in the program evaluated by Cloward (1967); but the primary focus was upon reading, structured SRA reading laboratories were used in the centers, and the tutors received separate instruction each week. The program developed by Glatter (1968) was the least structured of all the successful programs employing control groups, but even in this program there were specific times for meeting, and the primary focus appeared to be upon arithmetic computation. Given these guidelines, it is quite possible that the program described by Glatter would not have been successful if he had used criterion tests on reading comprehension or arithmetic problem solving instead of the arithmetic computation test which he employed.

The Hassinger and Via project involved a great deal of supervision of the tutoring process and a concentrated preservice training course for both the teacher-supervisors and the teen tutors in reading and reading materials, conducted by reading specialists. The Lundberg report

is the only fairly unstructured study to show results. However, interpretation of this study is confounded by the amount of self-initiation required by the experimental subjects in requesting tutoring, structuring their own time segments, and deciding their own meeting places and terminating points. Both the type of student and the student's own "structuring" attempts are variables unaccounted for in the study.

The Logan-Cache experiment concentrated heavily on specific skills and had a rigorous training and supervisory design for its tutors.

Although the successful tutoring programs all showed evidence of some form of structuring, most of the unsuccessful programs were unstructured and fairly unfocused.

The parent-aides in the program studied by Kirk (1966), or those who did not follow programed tutoring in the study by Ellson et al. (1968), had much greated freedom to select materials and activities. The anecdotal reports cited by Kirk suggest that a good deal of time was spent in discussion and meeting affective needs, and less time in reading tutoring. The program evaluated by Rosenshine and Furst (1969) was much less structured, and the section leaders of the different educational psychology classes differed widely in their conceptions of what constitutes a tutoring program. In the study by Weitzman (1965), in which the teacher's own tests were the criterion, it is quite possible that the tutors were unaware of the specific content covered on the teachers' tests.

The demonstration phase of the South-Western City School District teen tutorial project (1969) was also highly supervised and structured in terms of the child development components taught to the tutors.

However, within the kindergarten classroom, activities ranged from the tutor's "helping a child learn to use a scissors or hold a pencil" to seeing that the teacher's instructions are understood and to providing individual practice at whatever the kindergartner seemed to need at the moment. There were, thus, no programed or structured materials employed; nor were tutors expected to aid in an organized, step-by-step fashion in developing particular skills.

Time Devoted to Tutoring. There is no evidence of any optimum frequency of tutoring sessions. Tutoring programs which met once a week have been successful (Werth, 1967; Glatter, 1967) and unsuccessful (Cloward, 1967; Rosenshine and Furst, 1969). Indeed, in two separate years Kirk found no correlation between amount of time spent in tutoring and pupil achievement.

Optimal Groupings. Shaver (1969) reported no significant differences between one-to-one and one-to-three tutoring instruction in three replications. One-to-five ratios also seem to produce the same effects as smaller groupings. However, there are insufficient data on the one-to-five tutoring ratio to warrant any firm conclusions.

Characteristics of a Successful Tutor

The tutors in these studies have included parent-aides, college students, and low achieving high school students. Seemingly relevant characteristics such as age, experience, academic attainment, academic aptitude, or measures of attitude do not appear to be related to successful

or unsuccessful tutoring. Investigators who have made more detailed studies of tutor characteristics within a single investigation have not uncovered strong correlates of tutoring success. Glatter (1967) did not find any relationship between pupil achievement and the following characteristics of college tutors: knowledge of subject matter, self-concept, and level of anxiety. Cloward (1967) did not find any relationship between aptitude and achievement characteristics of the tutor and pupil achievement. Successful programs have been run with parent aides (Ellson et al., 1968), low-achieving high-school students (Cloward, 1967; Werth, 1967) and college students (Glatter, 1967); unsuccessful programs have also been run with parent-aides (Ellson, 1968; Kirk, 1966), low achieving high school students (Grannick, 1968), and college students (Rosenshine and Furst, 1969).

Tutoring in the Future

A cursory reading of this review might lead the educator to conclude that tutoring should be abandoned. That is a conclusion which is farthest from the minds of these reviewers, who have devoted literally thousands of hours to this report. The primary conclusion we wish to see drawn from the preceding pages is that tutoring should be expanded and not decreased. However, any expansion of tutoring should clearly concern itself with the following:

1. Concentrated efforts at more evaluative studies of tutoring. The number of objective assessments of tutoring reported in this monograph is rather small in proportion to the number of tutoring projects

which are being conducted. In preparing this review, we and our assistants searched through the ERIC collection, Educational Index, bibliographies of tutoring reports, Dissertation Abstracts; we have followed leads given to us by friends, and have called and written to investigators to see if they could direct us to further studies. Our results are meager, but the total number of studies reported here is larger than the number reported in any previous review.

As a result of our search, we have found 10 studies on cognitive achievement which utilize reasonable experimental design, three studies in which achievement data was collected but control groups were not used, and numerous programs which limited themselves to overall description but reported no achievement data. Overall, there was a negative relationship between the rigor of the design and the success of the program; the descriptive studies report much more "success" than those which employed control groups and statistical analysis of the results.

Our review revealed a dearth of published materials, and our findings are similar to those of Lundberg(1968), who reported that of the 33 school districts in California which made use of peer tutoring, not one had evaluated its program. More successful programs must be built from knowledge of results of past programs, and these results are not readily available.

2. More publicity for evaluative reports. Whether or not the reports are favorable to the tutoring project under consideration, reports must be made available. We had some real difficulties in locating reports

which we knew had been issued and which included negative findings. For some reason or other, these were "unavailable" from funding agencies. In one case, we were able to obtain a report only from the personal files of the original investigator.

Discussions of theoretical advantages of tutoring, with no attempts at evaluative practices, or "hiding" unfavorable conclusions, seem unwise. At best these practices lead to "horse and buggy" programs in a "space age," or to the expenditure of money with no tangible results. At worst, they tend to dupe the general public into believing that tutoring holds promise for far more than it is capable of delivering. This may lead to unnecessary disillusionment and bitterness about education and about the practice of tutoring in particular.

3. More efforts at replicating successful programs and program components. If, in fact, our educated "hunches" about the successes of the more structured programs are valid, further replications of studies involving highly structured programs would seem to be mandatory. This can be done, however, if researchers provide:

4. More information about the objectives of their programs, details of the tutor training, descriptions of the materials used in tutoring situation. All of these should be given with as many specific examples and actual materials as possible. Without this knowledge it is difficult to synthesize the results in any meaningful way, and it is almost impossible to replicate programs.

5. A clearer understanding and acceptance of the difference between objective and subjective criteria. Much of the preceeding discussion dealt with the great amount of subjective, anecdotal type of reporting associated with tutoring evaluations. The purpose of stressing the kind of data that is made available is not to demean the importance of observers' ratings, or of the perceptions of school personnel, parents, tutors, or the tutees themselves, about the experience. We are suggesting, however, that more rigorous efforts need to be undertaken to separate more clearly the two types of evaluation. Normally hard-nosed researchers have accepted a multitude of criteria for "success," and have agreed to continue some very costly projects on the basis of nebulous, or virtually nonexistent data. The "feelings" expressed by tutors, teachers, and principals (Rosenshine and Furst, 1969), anecdotal reports (Lippitt and Loman, 1965 and the fact that the schools want to continue a program (Hassinger and Via, 1969) are examples of reasons for continuing tutoring programs which are cited in the literature.

An interesting approach to building multiple criteria has been developed by the South-Western City School District in its use of a Profile of Evaluation based on both objective and subjective evaluations. Unfortunately, however, their acceptance of the efficacy of the demonstration phase of the teen tutorial program came almost exclusively from the anecdotal data in the profiles. In all cases where objective data and subjective data were available for the same hypothesis, the two were in conflict, and the subjective criteria were accepted.

It is hoped that more understanding of the interrelationships of different data collecting procedures will be developed. Hopefully, there

will also be attempts at building better measuring instruments, expecially in the affective domain.

At a minimum, it is suggested that those who believe that the effects of tutoring "cannot be measured by extremely careful in their publicity efforts in favor of the practice, and refrain from making claims that cannot be substantiated.

6. Longitudinal studies. Only one study was found in which careful, longitudinal followups have been done to assess the effects of tutoring after an elapsed time interval (Shaver, 1969). It is obvious that more work along these lines is needed.

7. More realistic expectations. It should be recognized that the results of the successful and unsuccessful tutoring programs reviewed here suggest that tutoring programs, even under the best of circumstances, will not achieve massive gains in a short period of time. There is no evidence here for the frequently voiced pronouncement that "turning kids on" or "treating them as individuals" will bring strong gains in both reading and arithmetic.

Bringing low achieving pupils to the implicit goal of "grade level" will take a long time, and directors and participants in tutoring programs should develop programs which will last from two to four years, and in which individual pupils will be kept not just for a set period of time, but until they reach and surpass the desired objectives. If we are to use other measures of "success," such as getting students up to their "potential" (Shaver, 1969), massive changes in school evaluation procedures need to be undertaken.

8. Focus on "achievement," or attaining objectives--regardless of how they are defined--rather than on a period of time. If the focus of tutoring is upon achievement of program objectives rather than on a period of time, then we might hope that future reports would be cast in a different form. In place of the current format in which an experimental group is compared to a control group for a set period of time, future reports might begin by stating the level of the children when they began the program and conclude not by stating whether they differed significantly from a control group after 10 weeks, but by stating how long it took to bring all the participating children to a desired level of mastery.

This desired level of mastery may be grade equivalent scores or may be "potential" scores or ther measures. The important element her is a more realistic view by the investigator of what may be accomplished. If grade equivalent scores are important, the time periods for these projects obviously needs to be increased. If mastery of other criteria is important, these should be clearly delineated, Only then will we have any idea of the time necessary for effective tutoring.

References

Bernstein, B. The after-school centers for disadvantaged pupils. New York: Center for Urban Education (ERIC ED 011 277), 1966.

Brookover, W. B., et al. The relationship of self-image to achievement in junior high school subjects. U.S.O.E. cooperative research report, project No. 845. East Lansing, Mich.: Michigan State University, Office of Research and Publications, 1962.

Brookover, W. B., et al. Improving academic achievement through students' self-concept enhancement. U.S.O.E. cooperative research project No. 1636. East Lansing, Mich.: Michigan State University, October 1965.

Carton, A. S. Corrective reading services for disadvantaged pupils in non-public regular schools. New York: Center for Urban Education.

City Board of Education, Logan, Utah. Tutorial instruction for underachieving readers and writers, 1968.

Cloward, R. D. Studies in tutoring. Journal of Experimental Education, 1967, 36, 14-25.

Deering, A. Fact-sheet: Homework helper program. New York: Board of Education, October 1968.

Eaverly, R. A. The 1967 summer tutoring project of youth serving youth (project "involve"). Philadelphia, Pa.: School District of Philadelphia, Office of Planning.

Eiseman, J. W., & Lippitt, P. Olders-youngers project evaluation. Ann Arbor, Mich.: University of Michigan, Institute for Social Research, 1966. (mimeo.)

Ellson, D. G. Report of results of tutorial reading project: 1966-1967. (mimeo.) Bloomington: Indiana. Dept. of Psychology, Univ. of Indiana.

Ellson, D. G., Harris, P., & Barber, L. A field test of programed and directed tutoring. Reading Research Quarterly, 1968, 3, 307-367.

Furst, N. F., Rosenshine, B., & Mattleman, M. S. Tutoring the low achieving urban child: A research report and review. Paper presented to the American Educational Research Association. Philadelphia, Pa.: Temple University, 1970.

Glatter, A. N. Significant variables of short-term tutoring. Doctoral dissertation (UM 68-9960). Detroit, Mich.: Wayne State University, 1967.

Goldberg, M. Methods and materials for educationally disadvantaged youth. In Passow, Goldberg, & Tannenbaum (Eds.), Education for the disadvantaged. New York: Holt, Rinehart and Winston, Inc., 1967.

Grannick, L. Youth tutoring youth: It worked. New York: National Commission on Resources for Youth, 36 West 44th Street, New York, N. Y. 10036.

Hamblin, J. A., & Hamblin, R. L. On teaching disadvantaged preschoolers to read: A successful experiment. American Educational Research Journal, 1972, 9, 209-216.

Hassinger, J., & Via, M. How much does a tutor learn through teaching reading? Journal of Secondary Education, 1969, 44, 42-44.

Hawkridge, D. G., Chalupsky, A. B., & Roberts, A. O. H. A study of selected exemplary programs for the education of disadvantaged children. Parts I and II. Palo Alto, Calif.: American Institutes for Research, 1968. (Part I: ERIC: ED 023 776; Part II: ERIC: ED 023 777)

Kirk, J. An evaluation of the school volunteer pilot program. Yellow Springs, Ohio: Community Service, Inc., 1966. (mimeo.)

Lippitt, P., & Lohman, J. E. Cross-age relationships: An educational resource. Children, 1965, 12, 113-117.

Lippitt, R., & Lippitt, P. Cross age helpers. National Education Association Journal, March 1968.

Lundberg, D. L. V. Some evaluations of tutoring by peers. Journal of Secondary Education, 1968, 43, 99-103.

McGraw, F. The use of 35 mm time-lapse photography in providing feedback to student teachers on the attending behaviors in their classrooms. Unpublished doctoral dissertation, Stanford University, 1965.

Niedermeyer, F. C., & Ellis, P. A. The development of a tutorial program for kindergarten reading instruction. Inglewood, Calif.: Southwest Regional Laboratory for Educational Research and Development, 1970.

Passow, A. H. Education in depressed areas. In Passow, Goldberg, & Tannenbaum (Eds.), Education for the disadvantaged. New York: Holt, Rinehart and Winston, Inc., 1967.

Rosenshine, B., & Furst, N. F. The effects of tutoring upon achievement and attention of under-achieving pupils. Paper presented to the American Educational Research Association, 1969. Urbana, Ill.: University of Illinois, Bureau of Educational Research.

Silver, A. Extending a helping hand educationally. Bakersfield, Calif.: Bakersfield College, Office of Institutional Research, 1967. (ERIC: ED 017 267, mimeo.)

Southwestern City School District, Grove City, Ohio. Teen tutorial program to meet the developmental needs of disadvantaged children in this and future generations, 1969.

Weitzman, D. L. The effect of tutoring on performance and motivation ratings in secondary school students. California Journal of Educational Research, May 1968, XVI, 3.

Werth, T. G. An assessment of the reciprocal effect of high school senior low achievers tutoring freshmen low achievers in English classes. Corvallis: Oregon State University, 1968. (UM 68-14, 822)