DOCUMENT RESUME

ED 239 220 CS 007 441

AUTHOR Meyer, Linda A.; And Others
TITLE Direct Instruction: A Project Follow Through Success Story. Technical Report No. 302.
INSTITUTION Bolt, Beranek and Newman, Inc., Cambridge, Mass.; Illinois Univ., Urbana. Center for the Study of Reading.
SPONS AGENCY National Inst. of Education (ED), Washington, DC.
PUB DATE Dec 83
CONTRACT 400-81-0030
NOTE 34p.
PUB TYPE Reports - Evaluative/Feasibility (142)

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS Basic Skills; *Disadvantaged Youth; Elementary Education; Faculty Development; Inservice Teacher Education; Preservice Teacher Education; *Program Evaluation; *Reading Achievement; Reading Improvement; *Reading Instruction; *Reading Programs; *Reading Research; Teaching Methods
IDENTIFIERS Direct Instruction; Project Follow Through

ABSTRACT

The Follow Through Project at New York City's P.S. 137 has demonstrated that even students from highly disadvantaged areas can match the academic accomplishments of their middle class peers. Begun in 1968, the project involved 12 classrooms at the kindergarten through grade 3 levels and used a direct instruction model featuring (1) a consistent focus on academic objectives; (2) high allocations of time to small group instruction in reading, language, and math; (3) a carefully sequenced Distar curriculum; (4) preservice and inservice teacher training; and (5) a comprehensive system for monitoring both the rate of students' progress and their mastery of the materials covered. Reasons for the project's success included the administration's commitment, parental support, and the continuity and structure that the instructional model gave to a system with a high staff turnover rate. Evaluations conducted by a variety of sources indicated that P.S. 137 students' achievement scores compared favorably with students from similar backgrounds. In addition, the local school district's review of long term effects showed that Follow Through students maintained mean scores at or above grade level in grades 4 and 5. (Tables summarizing evaluation findings are included.) (MM)

CENTER FOR THE STUDY OF READING

Technical Report No. 302

DIRECT INSTRUCTION: A PROJECT FOLLOW THROUGH SUCCESS STORY

Linda A. Meyer
University of Illinois at Urbana-Champaign
Russell M. Gersten
University of Oregon
Joan Gutkin
New York University

December 1983

University of Illinois
at Urbana-Champaign
51 Gerty Drive
Champaign, Illinois 61820

Bolt Beranek and Newman Inc.
50 Moulton Street
Cambridge, Massachusetts 02238

The research reported herein was supported in part by the National Institute of Education under Contract No. NIE-400-81-0030. A modified version of this paper appears in Elementary School Journal, 1983. Ronald P. Toby, three anonymous reviewers and Thomas L. Good made suggestions that strengthened this manuscript.

Direct Instruction: A Project Follow Through Success Story

A variety of large-scale field studies have consistently shown that the average reading and math achievement levels of low-income, minority students usually are at the 20th to 28th percentiles by the end of third grade, virtually a year below grade level (Ozenne, et al., 1976; Molitor, Watkin, Napier, & Proper, 1977; National Assessment of Educational Progress, 1979). The educational changes--both institutional and instructional--necessary to improve this situation have been difficult to accomplish (Stebbins, St. Pierre, Proper, Anderson, & Cerva, 1977) particularly in inner-city schools (Cohen, Koehler, Datta, & Timpane, 1980).

In 1968, the U.S. Office of Education initiated a comprehensive program called Project Follow Through for economically disadvantaged children in the primary grades in 180 communities. Unlike Headstart or the subsequent Elementary and Secondary Education Act Title 1 programs, each local Follow Through project was aligned with an outside sponsor: a university, educational laboratory, or state department of education. This alignment represented a unique, innovative educational model. A wide array of instructional approaches were included in Follow Through, ranging from open classroom models to cognitive models based on the theories of Piaget, to highly structured programs utilizing principles of contemporary learning

theory. The sponsor was responsible for designing and implementing a comprehensive educational program in each project. An equally wide range of communities as included in Follow Through--from rural communities like Flippin, Arkansas to large urban areas such as New York City and Philadelphia.

This paper describes one of the nine Follow Through projects in New York City, Project P.S. 137, which was conducted in the Ocean Hill-Brownsville section. The project involved twelve classrooms in one school, three at each grade level from kindergarten through third grade. The program in fourth and fifth grades was a traditional New York City curriculum. The P.S. 137 Project was affiliated with the Direct Instruction Model from the University of Oregon, a highly structured educational model. The other eight New York City Follow Through projects were aligned with other models.

A study was conducted by the U.S. Office of Education to explore the effects of the various educational models in two of the country's largest urban areas, New York City and Philadelphia. The authors of the Abt Report (Stebbins, et al., 1977) identified several characteristics of large cities that made successful delivery of any program particularly difficult--"high teacher turnover, teacher strikes, formal negotiations over teacher contracts, and the bureaucracy generally associated with large school systems" (Stebbins, et al., 1977, vol. IV-A, p. 150). They viewed the big cities as a "test of the educational

model's ability to adapt to often adverse conditions, a test which appears difficult to pass" (vol. IVA, p. 148).

Within this context Stebbins and colleagues compared the results of eleven Follow Through projects in the two cities. Only one of these eleven projects--the P.S. 137 Direct Instruction Model--showed consistent, significant positive effects in all academic areas--both basic skill areas (e.g., arithmetic computation) and higher order cognitive skills (e.g., reading comprehension, math problem solving). The same project was recently approved for national validation as a successful program by review of the National Institute of Education (Gersten, Meyer, & Gutkin, 1981).

In light of the consistently disappointing educational results in inner-city settings like Ocean Hill-Brownsville, it makes sense to describe the critical variables that constitute the program. Although on the basis of summative evaluation data it is impossible to isolate the factor or factors that led to success, identifying the components of the program may assist schools in comparable settings to develop programs which will achieve similar results.

Ocean Hill-Brownsville

The Ocean Hill-Brownsville section of Brooklyn has long been recognized as one of the most economically and educationally disadvantaged areas in the United States (New York Times, February 4, 1980). The area suffers from high unemployment,

reliance upon welfare, low levels of literacy, substandard housing, insufficient living space, and racial isolation.

According to the 1970 census, almost 75% of the adults (16 years old and over) living in Ocean Hill-Brownsville have completed less than eight years of schooling. Pupils in Ocean Hill-Brownsville have for many years achieved the lowest reading scores of all the 32 school districts in New York City. Ninety-nine point nine percent of the students are from minority backgrounds. Student turnover in the district is estimated at about 40% annually (New York Times, February 4, 1980).

The well-publicized conflict between the administrators and staff at P.S. 137 and the United Federation of Teachers in 1966-1968 over the issue of community control irreversibly politicized the Ocean Hill-Brownsville district, and P.S. 137 in particular (Mayer, 1969). Related activities got parents involved in schools in such a way not found in any other poverty-ridden area. Parents learned how to use power; some used this power to have P.S. 137 chosen as one of the schools in the national Follow Through program, and to select the structured Direct Instruction model. Parent support has kept the program going for 14 years, despite cuts due to the New York City budget crisis of 1975 and subsequent reductions in Federal funding. During the years of the budget crisis, many experienced staff members (teachers and paraprofessionals) were transferred or laid off. Between 1968 and 1981, the project at P.S. 137 had a high turnover rate, with over fifty teachers, five principals, and six Follow Through

Directors staffing its twelve classrooms. Despite these conditions, the model has endured, adapting to the training needs of new teaching personnel and administrators.

The Direct Instruction Model

The Direct Instruction model has the following components:

(1) a consistent focus on academic objectives;

(2) high allocations of time to small-group instruction in reading, language, and math;

(3) the tight carefully sequenced Distar curriculum, which includes a task analysis of all skills and cognitive operations and numerous opportunities for review and practice of recently learned skills;

(4) ongoing in-service and pre-service training which offers concrete, "hands on" solutions to prbiems arising in the classroom;

(5) a comprehensive system for monitoring both the rate at which students progress through the curriculum and their mastery of the material covered.

More complete descriptions of the curriculum and the philosophy of instruction are presented elsewhere (e.g., Becker, Engelmann, Carnine, & Rhine, 1981; Becker & Carnine, 1981; Carnine & Gersten, 1982). In this paper we will describe the monitoring, administrative, and supervisory elements of the model. We believe these elements are of great relevance for curriculum systems other than Distar, particularly when active

teaching (Good & Grouws, 1979) or other direct instruction approaches (e.g., Stallings, 1980) are being implemented.

Sponsorship: The project manager. Central to the success of the Follow Through project at P.S. 137 was the relationship of the school staff to the Project Manager appointed by the University of Oregon (the sponsor). The Project Manager is more than a consultant; he/she is responsible for transmitting the model to the school. The manager spends between 20 and 40 days a year at the school; at least half that time is spent conducting in-service training, and meeting with parents and administrators. The manager plays an active role in the development of classroom schedules, the monitoring of teacher and student performance, and the assignment of students and staff.

Curriculum materials. The Distar curriculum differs from other curriculum programs in that it provides a teacher presentation for each lesson. The manual indicates not only the general manner of the presentation, but also the exact wording to be used by the teacher for the lesson. There are sequenced lessons for reading, language, and arithmetic. The local staff supplements these materials with a series of written comprehension questions to a linguistic reading series, and with a basal reading program in third grade. They teach the regular New York City curriculum in other subject areas such as social studies, science, and art.

Student materials such as readers, workbooks, and "take-home sheets" are coordinated with the teacher presentation books.

Generally, there are three homogeneous (based on ability) instructional groups of six to ten students in each class. As each group completes a level, subsequent levels are introduced, regardless of grade.

The careful specification of teacher and student behaviors in the model means that a supervisor should be able to walk into any classroom, look at the lesson number (e.g., Arithmetic II, Lesson 15) and have a clear idea of what should be happening. Furthermore, the careful sequencing of the lessons makes training easier, and these materials guarantee more consistency from one teacher to another. Also, because the lessons are scripted, the children know almost as well what to expect from their substitute teachers as regular teachers.

Allocated time. Approximately three hours per day, or a little less than sixty percent of the available school day, is allocated to instruction in reading, math, and language (one hour per subject). Each student group receives thirty minutes of teacher-directed instruction in each of these areas, and spends thirty minutes working independently to complete assignments that reinforce and provide practice on skills presented during the teacher-directed activities.

Monitoring instruction. To judge and criticize teacher performance without offering suggestions on how that performance can be improved seems wrong. In contrast, the Direct Instruction Model examines everything from textbooks and critical teaching

behaviors to placement decisions and procedures for assessment. The model specifies in detail what can be done to improve consistent errors (e.g., a child who has problems identifying the main idea of a story, a poorly motivated reading group) and provides precise guidance and feedback on the implementation of the solution in the classroom. The desired teacher behavior can be modeled by the supervisor in the classroom (Becker, Engelmann, Carnine, & Rhine, 1981; Carnine, & Gersten, 1982).

Monitoring both the quality and quantity of instruction is a key element of success. Since the lessons in each subject are numerically sequenced, from 1-160, it is not difficult to monitor the quantity of instruction.

At the beginning of each year, the Project Manager meets with each classroom teacher and paraprofessional, and they determine the number of lessons to be completed by each group in the class for the school year. Average-ability groups are expected to complete one lesson each day in each subject area; adjustments are made for higher and lower performing groups. Every two weeks teachers and paraprofessionals submit a Lesson Progress Report Form. The results go to the Project Manager as well as the principal and local administrator. Every three months the progress for each group is calculated and the teacher and paraprofessional meet with the Project Manager to discuss each group's progress and develop strategies for acceleration or review.

Figure 1 demonstrates a completed Lesson Progress Report Form for a two-week period. It shows what lesson each group is on and how many lessons they completed in the preceding two weeks. This report was submitted on the 107th day of school. Group 1 has gained twelve lessons in the last ten day period in reading, while the second group gained 10 lessons, and the third group 8 lessons. Each group made comparable progress in arithmetic.

Insert Figure 1 about here.

Quantity of instruction without corresponding quality is meaningless. If the number of lessons was the only measure of instructional effectiveness, teachers might be inclined to "turn the pages," i.e., to complete lessons regardless of how students were performing. In fact, we have observed many novice teachers doing this. To avoid this danger, criterion-referenced tests are administered by a trained tester (someone other than the classroom teacher). This is a far more objective system than teacher-administered tests. Students are tested in either reading, language, or math every three weeks. With this schedule student performance is monitored every three weeks in one subject area and every nine weeks in all three areas. Testing in the lower grades is done on an individual basis. By third grade, the tests are often group administered.

Sample scores from these tests are reported in Figure 2. These results are from the middle group, on lesson 82, about half way through the first level of the reading program. Note that there are seven children in the group. Two children scored 86%, two scored 72%, one scored 57%, and one 43%. The seventh child was absent. The scores in the extreme right column are for individuals. The percentage scores at the bottom of each column are for the test items in this segment of the arithmetic test. There were seven items. There are three items scored at 93%, and one at 100%, two at 33%, and one at 50%.

Insert Figure 2 about here.

These data show that two members of the group are doing well, two are having problems, and two are in serious trouble. However, overall mastery is high for the first four items, and low on the last three. Thus remediation is an individual problem for the first four items, and a common group problem with the last three items. The teacher will focus on the last three skills with the group; and the Project Manager will observe the teacher, to see if the difficulty is in the teacher's presentation.

Teacher training. The Project Manager conducts training on either a one-to-one basis or in groups outside the classrooms, depending upon the number of staff with common problems. Though these sessions sometimes include explanations of the rationale

behind certain teaching procedures, the major emphasis is on practicing various techniques that the staff will use.

Once teachers have undergone initial training, they receive weekly technical assistance in their classrooms. The primary mode of supervision is direct observations of teachers in their classrooms by the Project Manager or a teacher trainer. After each observation, teachers receive written feedback. Often the teacher trainer will actually teach the class for a five or ten minute segment in order to demonstrate a new approach for motivating a daydreaming student, or providing constructive feedback to children who have errors. It is therefore essential that the teacher trainer and Project Manager be skilled teachers. Supervisors often give teachers weekly assignments (i.e., practice on new skills and techniques).

The Direct Instruction model is a comprehensive and complex one. In a sense, it covers all the bases by developing close lines of sponsor supervision with ongoing preservice and inservice training. Managers and consultants demonstrate for teachers and paraprofessionals, actually presenting models of what instruction should look like. The materials articulate what, when, and how the teachers should teach, and how the students should perform. And, to assure that all of this is happening, lesson progress and performance reports go to all involved parties.

A high degree of structuring, attention to detail, and the high level of parent support in a very real sense constitute the "program," and together they have produced the student achievement gains that show that the program is significantly effective for this student population.

Evaluation of the P.S. 137 Project

So many evaluations of this Follow Through project have been conducted that the first question to arise is which data to present. The program began with students entering kindergarten in 1968. Data are presented from 1973 (when the second cohort of students reached third grade) through 1981. Results of three separate evaluations are presented.

Independent evaluation by Abt Associates. A major independent evaluation was conducted for the U.S. Office of Education by Abt Associates and the Stanford Research Institute (SRI). This evaluation intensively examined two cohorts of children, those beginning Follow Through in 1970 and completing third grade in 1972 (Cohort II), and those beginning kindergarten in 1971 and concluding third grade in 1974 (Cohort III). The Abt Report studied several sites and assigned comparison groups for the nine largest sponsors. Unfortunately, since Follow Through served the neediest students in a community, students in the comparison groups tended to be somewhat less disadvantaged than the Follow Through students (House, Glass, McLean, Walker, 1978; Stebbins, et al., 1977). New York City had one Direct

Instruction Follow Through project and SRI thus selected a comparison school in New York City.

Upon entry into kindergarten, children in both Follow Through classrooms and comparison (Non Follow Through) classrooms were tested on the Wide Range Achievement Test (Jastak & Jastak, 1966). In addition, demographic information (sex, family income, mother's education, ethnicity, home language) was collected. All testing was done by SRI. In the spring of the third grade, students who had been in Follow Through for the full four years (or had remained in the Non Follow Through school for four full years) were given all subtests of the Metropolitan Achievement Test (Durost, Bixler, Wrishstone, Prescott, & Balow, 1970).

Using analysis of covariance, scores for the Follow Through students were compared to scores of (a) students in the local comparison (NFT) sample and (b) a "pooled comparison" sample of 6,000 low SES students.[1] Covariates included SES, pretest scores, ethnicity, and home language. The latter comparison should be less biased by covariance adjustments because the size of the comparison sample was so large and so many different communities were sampled that any idiosyncracies or inequities in local sampling would be minimized.

Table 1 presents the results of the Abt evaluation of achievement for Cohorts II and III at P.S. 137 and a New York comparison group. Descriptive statistics are presented for students on all subtests of the Metropolitan Achievement Test. These comparisons were made to offset any bias due to the local

comparison group being a bit less (or more) disadvantaged than Follow Through, or the program in the local comparison schools being a bit better (or worse) than existing educational practices for low income students (Stebbins, et al., 1977, vol. IV-A). The mean raw scores have been converted to percentiles to give the reader a sense of how Follow Through students compare to the norm sample of the MAT. Table 1 indicates that Follow Through students performed at or near the national median in all measures. For example, Cohort II is at the 54th percentile in Total Reading and the 56th percentile in Total Math. Cohort III is also at the 54th percentile in Reading and even higher (66th percentile) in Math. The column on the extreme right presents the magnitude of the covariance-adjusted treatment effect in pooled standard deviation units. Generally, any effect of .25 standard deviation units or more is considered educationally significant.

Insert Table 1 about here.

Statistically significant positive effects are found when comparisons are made with the pooled group rather than with the somewhat less disadvantaged local comparison school (see House, et al., 1978 or Carnine & Gersten, 1982 for a more thorough discussion of the covariance analyses). Any magnitude of the treatment effect larger than one-fourth standard deviation is

considered educationally significant (Horst, Talmadge, & Wood, 1975).

University of Oregon and New York City testing programs. The quasi-experimental design used by Abt controls for most of the traditional threats to internal validity such as maturation, reactivity to testing, and history. However, as Cook and Campbell (1979) state, there will always be potential flaws in any quasi-experimental field study. Thus, the only way to demonstrate effectiveness convincingly is to show replicability across time (i.e., across cohorts of children).

Table 2 shows end of third grade achievement scores for Cohorts I through IX in reading. The sample includes only low-income children in the Follow Through program for four full years. Cohorts II through VI were tested on the MAT under the supervision of the University of Oregon. Cohorts VI through IX were tested on the SAT under the supervision of New York City Follow Through. The Anchor Study (Loret, Seder, Bianchini, & Vale, 1974) demonstrated that the MAT and SAT are reasonably comparable. For purposes of comparison, the mean Total Reading score for comparable Non Follow Through children in large urban centers in the Northeast, gathered by Abt in 1974-1975, is presented. These figures were corroborated by subsequent research (Ozenne, et al., 1976), and appear to be a reasonable comparison standard for the children in P.S. 137.

Insert Table 2 about here.

The drop in reading for Cohorts VI and VII was likely due to the budget crisis in New York City which began in 1975-76. Reduced budgets led to fewer teachers and paraprofessionals in the city, and to the loss of a teacher trainer and a family worker position at P.S. 137. Citywide, tenured teachers and paraprofessionals were often reassigned to schools on the basis of seniority, thus causing a great deal of staff mobility. In addition, less money was available for instructional materials or stipends for parents to come to the school for training or to tutor in the classrooms. Cohorts VIII and IX, with percentile ranks of 46 and 47, show a marked increase in reading scores as a period of stability again emerged. Even with limited services and constant staff mobility, the P.S. 137 children still performed significantly higher in reading than inner-city students in the Northeast region.

Long-term effects: Evaluation conducted by the local district. Data were also collected by the district, which compared the performance of Follow Through students who completed the program with that of other students in the school district who received traditional educational programs. Follow Through scores are compared to the District's scores because of the demographic similarity of P.S. 137 to other (Non Follow Through) schools in the district. These data are of particular interest

because they allow for an examination of how students perform after they have completed the Follow Through program. This testing consisted of the administration of the MAT in 1974; the SAT in 1975 and 1976; and the CTBS in 1977. The third, fourth, and fifth grade scores of all graduates of Follow Through Cohorts III and IV were traced. All scores are reported as mean grade equivalents, rather than standard scores. Mean Grade Equivalents (GE's) are used in these followup analyses only, since the New York City testing program reported scores in GE's only. Though grade equivalent scores are not an interval score, and therefore are not as precise as the standard score units used in the other analyses, they should offer reasonably good estimation of the effects. It is unlikely that the use of grade equivalents would systematically bias the test results for within-grade comparisons which we present. Table 3 shows these comparisons. Not only did the Follow Through students maintain mean scores at or above grade level in Grades four and five, but they scored significantly higher than the remainder of the students in the district ($p < .05$). Thus, there is evidence that the positive effects of Follow Through are maintained in the intermediate grades.

Insert Table 3 about here.

Conclusions

Mean performance for comparable disadvantaged inner city students on standardized achievement tests in third grade Reading and Math typically is between the 24th and 30th percentiles. The students in the program at P.S. 137 have consistently surpassed these levels. Given the Direct Instruction model as it has been described, what are the implications for districts with similar conditions?

In her secondary analyses of the Follow Through evaluation, Kennedy (1978) argued that perhaps the "assertive," non-adaptive tactics used by the University of Oregon to insure that its model was implemented as conceived led to its success in New York City, when none of the other Follow Through approaches succeeded in the large cities. Edmonds (1979) noted that in the effective inner-city schools he observed, principals adopted the same assertive role--insuring adequate time was spent each day in reading, instilling high expectations for all students to succeed, actively monitoring progress in reading. These are quite similar to the roles adopted by the University of Oregon's Project Manager.

Especially in school settings with high student and teacher turnover, it appears that a clearly specified, well articulated program has greater potential for continuity and success than models that are not well articulated. The Direct Instruction programs are less dependent than other programs upon the unique contributions of specific people. Once teacher trainers becom-

experienced, they can teach the basic techniques to new teachers in a matter of hours. This has been important in inner city areas where there is usually high teacher turnover. Others have also found that concrete, well articulated models of teaching can lead to improvement in achievement of pupils in urban schools (e.g., Good & Grouws, 1979; Stallings, 1980).

Traditionally, little monitoring of instruction or student progress has occurred in large school districts. The administrators and staff at P.S. 137 feel that the monitoring provided by criterion-referenced tests and the analysis of progress through the curriculum helped everyone to know how students were performing. This system of checks and balances could be implemented in other settings, with other curriculum materials.

The major difference between in-service training as it is defined in D.I. Model and that commonly provided in most districts is the continuity and consistency that characterize the D.I. Model, from pre-service to in-service training sessions to classroom observations and demonstrations. Berman and McLaughlin (1975) found that such concrete technical assistance to teachers was one of the leading factors in successful educational changes. Administrators or supervisors in other settings could follow this model and spend more time in the classrooms teaching model lessons and working with teachers and aides. Similar short-term attempts have been highly successful in studies by Stallings

(1980), Ebmeier and Good (1979), and Anderson, Evertson, and Brophy (1979). Teachers would probably welcome a break in the isolation that most of them experience if they received useful feedback and demonstrations rather than observations and brief follow-up conferences. Working with teachers in the classroom (rather than supervising teachers) shifts attention to student performance.

The parents of P.S. 137 selected the program for their children. A small group of vocal parents have supported the program strongly in a neighborhood where apathy is much more common. These parents have worked with the sponsor, attended parent training, and volunteered in classrooms. They back the program at the district and Central Board levels. They have been the strongest advocates of the model. They have helped to keep it alive. In his analysis of the data collected from parents at 16 Direct Instruction sites, Haney (1978) found that P.S. 137 parents disagreed with the statement that, "there is not much parents can do about changing the educational situation in their community." These parents viewed schools as helpful not only to their children, but also to themselves, particularly in terms of learning about teaching, learning how to help with their children's work, and understanding how their children learn. Of greater importance, parents affiliated with Direct Instruction, more frequently than other groups of Follow Through parents, felt that school had appreciably helped their children academically.

The Follow Through program at P.S. 137 shows the benefits of sponsorship in keeping a constant educational approach. Ongoing staff development and monitoring systems that are program-specific are also integral ingredients for success, as are teaching materials and techniques. The program proves that even students from highly disadvantaged areas, who have to overcome multiple handicaps both at home and in the school, can match the academic accomplishments of their middle-class peers. In a time of increased emphasis on basic skills and academic accomplishment, the Follow Through Program can proudly say it discovered these goals 12 years ago and has been meeting them ever since.

References

Anderson, L., Evertson, C., & Brophy, J. (1979). An experimental study of effective teaching in first grade reading groups. Elementary School Journal, 79, 193-223.

Becker, W. C., & Carnine, D. W. (1981). Direct instruction: A behaviorally-based model for comprehensive educational intervention with the disadvantaged. In A. Kazdin, & B. Lahey (Eds.), Contributions of behavior modification to education. Hillsdale, NJ: Erlbaum.

Becker, W. C., Engelmann, S., Carnine, D., & Rhine, R. (1981). Direct instruction models. In R. Rhine (Ed.), Encouraging change in American schools: A decade of experimentation. New York: Academic Press.

Berman, P., & McLaughlin, M. (1975). Federal programs supporting educational change, Vol. IV: The findings in review. Santa Monica, CA: Rand Corp.

Carnine, D., & Gersten, R. (1982). Improving reading instruction: A case study of Direct Instruction with some thoughts on the realities and logistics of educational change. Paper presented at Research Base for Literacy, a conference, organized for the Center for the Study of Reading, and sponsored by the Hegeler Institute, the Johnson Foundation, the Exon Education Foundation, the University of Illinois Education Fund, the Monsanto Fund, and the National Institute of Education, Racine, Wisconsin, March 12-14.

Cohen, M., Koehler, V., Datta, L., & Timpane, M. (1980). Instructionally effective schools: Research area plan. Unpublished manuscript. Department of Education, National Institute of Education.

Cook, T., & Campbell, D. T. (1979). Quasi-experimentation: Design and analysis issues for field settings. Chicago: Rand McNally.

Durost, W., Bixler, H., Wrightstone, J., Prescott, G., & Balow, I. (1970). Metropolitan Achievement Tests. New York: Harcourt, Brace, Javanovich.

Ebmeier, H., & Good, T. (1979). The effects of instructing teachers about good teaching on the mathematics achievment of fourth grade students. American Educational Research Journal, 16, 1-16.

Edmonds, R. Effective schools for the urban poor. (1979). Educational Leadership, 37, 15-24.

Gersten, R., Meyer, L., & Gutkin, J. (1981). Evaluation of P.S. 137 (New York City) follow through. Final report submitted to the joint dissemination review panel. Washington, D.C.: Department of Education.

Good, T., & Grouws, D. (1979). The Missouri mathematics effectiveness project. Journal of Educational Psychology, 71, 355-362.

Haney, W. (1977). A technical history of the national follow through evaluation. Cambridge, MA: Huron Institute.

Horst, D. P., Tallmadge, G. K., & Wood, C. T. (1975). A practical guide to measuring project impact on student achievement. Monograph No. 1 on Evaluation in Education. Washington, D.C.: U.S. Government Printing Office.

House, E., Glass, G., McLean, L. D., & Walker, D. (1978). No simple answer: Critique of the follow through evaluation. Harvard Educational Review, 48, 128-160.

Jastak, T., & Jastak, B. (1966). The Wide Range Achievement Test. Wilmington, DE: Guidance Associates.

Kennedy, M. (1978). Findings from the follow through planned variation study. Educational Researcher, 7, 3-11.

Loret, P. G., Seder, A., Bianchini, J. C., & Vale, C. A. (1974). Anchor Test Study: Equivalence and Norms Tables for Selected Reading Achievement Tests (Grades 4, 5, 6). Office of Education Report No. 74-305. Washington, DC: U.S. Government Printing Office.

Mayer, M. (1969). The teacher's strike: New York, 1968. New York: Harper & Row.

Moliter, J., Watkin, N., Napior, D., & Proper, E. C. (1977). Education as experimentation--the non-follow through study. Cambridge, MA: Abt. National Assessment of Educational Progress (NAEP). (1979). Mathematical knowledge and skills. (ERIC Document Reproduction Service No. 176 964)

New York Times. February 4, 1980.

Ozenne, D., et al. (1976). United States Office of Education. Annual evaluation report of programs administered by the U.S. Department of Education, FY 1975. Washington, D.C.: Capitol Publications, Educational Resources Division.

Stallings, J. (1980). Allocated academic learning time revisited, or beyond time on task. Educational Researcher, 9, 11-16.

Stebbins, B. (Ed.). (1976). Education as experimentation: A planned variation model, Vol. III. Cambridge, MA: Abt.

Stebbins, L. B., St. Pierre, R. G., Proper, E. C., Anderson, R. B., & Cerva, T. R. (1977). Education as experimentation: A planned variation model, Volumes IV A-D. An evaluation of follow through. Cambridge, MA: Abt.

Footnote

[1]There were eight other Follow Through projects in New York City representing seven other educational models. Each of these Follow Through projects was compared (statistically) to the New York City comparison group. Stebbins, et al. (1977) also compared the nine New York City projects to each other, concluding that the big cities could be the best test of a model's effectiveness. They concluded that only P.S. 137 had passed the test.

Table 1

The Abt Evaluation Summary of End-of-Third Grade Achievement on the Metropolitan Achievement Test

Outcome Measure	P.S. 137 Mean Raw	FT %ile	NYC Mean Raw	Comparison Group %ile	Magnitude of Effect in Pooled SD Units: Pooled Comp. Mag.	Magnitude of Effect in Pooled SD Units: Local Comp. Mag.
Cohort II						
TOTAL MATH	67.10	56th	59.05	44th	.58**	.99
TOTAL RDG.	55.65	54th	49.03	42nd	.38**	.01
Language	27.68	55th	17.72	28th	1.1 **	.91**
Cohort III						
	Median Stand.	%ile	Median Stand.	%ile		
TOTAL MATH	75.4[c]	66th	64.9[c]	32nd	1.08*	.82*
TOTAL RDG.	60.6[c]	48th	58.8[c]	42nd	.23	-.11
Language	76.1[c]	68th	59.3[c]	23rd	1.51*	1.36**

*p < .05
**p < .01

[a]The Pooled Comparison Group was based on all Non-Follow Through students.
[c]Median standard scores

Table 2

Achievement Test Data at the End of Third Grade from P.S. 137 for Cohorts I Through IX and a Comparison Group

	MAT Total Reading					SAT Total Reading				Non-Follow Through
Cohort	I	II	III	IV	V	VI	VII	VIII	IX	Northeast
Year	1973	1974	1975	1976	1977	1978	1979	1980	1981	Large City[b]
Percentile Equivalent	36th	52nd	42nd	46th	40th	33rd	36th	46th	47th	28th
Mean Standard Score	57	61.6	58.8	60.1	58	40.5[a]	42.3[a]	47.4[a]	48.1[a]	54.2
Standard Deviation	5.8	9.6	7.4	9.4	8.8	13.5	12.6	15.9	11.4	11.4
Sample Size(N)	31	36	38	46	36	31	41	32	19	688

[a] Mean Normal Curve Equivalent

[b] Mean score for all comparison (NFT) children in 2 large cities (New York and Philadelphia) for two cohorts of children (1973-1974). This figure corresponds to subsequent data collected by Ozenne et al., (1976) and NAEP (1979).

Table 3

Comparison of the Total Reading Scores for Follow Through Students in P.S. 137 and All Students in District 23, New York City

	Grade 3 - 1974		Grade 4 - 1975		Grade 5 - 1976	
	Mean G.E.	N	Mean G.E.	N	Mean G.E.	N
P.S. 137 Follow Through, Cohort III	3.72	34	5.19	31	7.63	26
District 23	3.0	**	3.80	1816	5.3	1798
	Grade 3 - 1975		Grade 4 - 1976		Grade 5 - 1977	
	Mean G.E.	N	Mean G.E.	N	Mean G.E.	N
P.S. 137 Follow Through, Cohort IV	4.02	37	4.72	30	5.77	27
District 23	3.1	1877	3.8	1824	4.6	1547

*Scores for all children are taken from the NYC Testing Program: MAT/1974; SAT/1975; SAT/1976; CTBS/1977.

**Unavailable

Figure 1

Lesson Progress Report Form

Site P.S. 137 Date April 3, 1981 School Day 107

Teacher		Reading			Arithmetic			Language		
	Group	Level	Day	Gain	Level	Day	Gain	Level	Day	Gain
	1	11	118	12	1	169	12	1	111	12
Allen	2	1	92	10	1	127	10	1	92	10
	3	1	74	8	1	74	8	1	132	8

Figure 2

GROUP SUMMARY FORM

Circle one:
(Reading I), II, III
Arithmetic I, II, III
Language I, II, III

Test Section 8
Lesson Number 82
Group II
Teacher Allen
Date 12/7/80

Names	Items 1	2	3	4	5	6	7	8	9	10	Percent Passed
1. Dave	+	+	+	+	-	+	+				86%
2. Sharon	+	+	+	+	-	-	+				72%
3. Bob	+	+	+	+	+	-	-				72%
4. Jane	+	+	+	+	-	+	+				86%
5. Bob	+	+	+	-	-	-	-				57%
6. Steve	+	-	-	+	+	-	-				43%
7. Marlene	Absent----------										
8.											
9.											
10.											
Percent passed	100	93	93	93	33	33	50				