8408997

Johnson, Juanita Jamison

EVALUATION OF THE READING PROGRAM OF THE GUILFORD COUNTY
SCHOOL SYSTEM

The University of North Carolina at Greensboro   ED.D. 1983

University
  Microfilms
International 300 N. Zeeb Road, Ann Arbor, MI 48106

EVALUATION OF THE READING PROGRAM

OF THE GUILFORD COUNTY

SCHOOL SYSTEM

by

Juanita Jamison Johnson

A Dissertation Submitted to
the Faculty of the Graduate School at
The University of North Carolina at Greensboro
in Partial Fulfillment
of the Requirements for the Degree
Doctor of Education

Greensboro
1983

Approved by

Dissertation Adviser

APPROVAL PAGE

    This dissertation has been approved by the following

committee of the Faculty of the Graduate School at the University

of North Carolina at Greensboro.

        Dissertation Adviser _____

        Committee Members _____

                          _____

                          _____

_____
Date of Acceptance by Committee

_____
Date of Final Oral Examination

ii

JOHNSON, JUANITA JAMISON. Evaluation of the Reading Program of
the Guilford County School System. (1983)
Directed by: Dr. Carol K. Tittle, Pp. 175

The purpose of this study was to determine the impact of

curricular intervention on the reading program of a local school

district. The study was based on the following assumptions: that

evaluation is part of a total curriculum improvement process; that

principals, teachers, parents and students are among the most

important audiences for the evaluation report and should therefore

be involved in the evaluation study; and that metaevaluation

techniques can be used to improve the validity, reliability and

objectivity of local evaluation studies.

Selected literature related to major evaluation models, to

the evaluation of reading programs, and to metaevaluation was

reviewed. Pretest and posttest survey data were analyzed to test

hypotheses relative to a change in the description of the program

and the satisfaction of principals, teachers, parents, and students

with the program. The reading achievement scores of third and

sixth graders in 1979 were compared to 1982 scores for third

and sixth graders to test an hypothesis related to student's

reading achievement. A metaevaluation study was conducted to

validate that determined effects were due to the curricular

intervention and not to other threats to internal validity.

The results of the study indicated that teacher's knowledge

of the school system's curricular guide had increased, the record

keeping of reading skills was more effective, use of the system's

curricular guide as a source of reading program goals had

increased, student achievement in reading had increased, and

teachers were more satisfied with the reading program.

Major conclusions drawn from the study were that a variety of

evaluation models exist and are being used to evaluate reading

programs; the purposes of the evaluation, the information that

is required and the audiences to be served help to determine

the models and/or combination of evaluation models that are

utilized; evaluations at the local level can be used to determine

program effects; and metaevaluation techniques can be used to

improve the validity, reliability, and objectivity of local

evaluation studies.

TABLE OF CONTENTS

Table of Contents (continued)

LIST OF TABLES

LIST OF TABLES (continued)

LIST OF FIGURES

CHAPTER I

INTRODUCTION

Program evaluation has become a key concept and an important operation in education today. This importance stems from several factors: the public demand for more accountability, agencies who want to know if funds are being used efficaciously, program partici- pants who desire feedback on their performance, and decision makers who are responsible for program effectiveness. Educators carry out these evaluative activities under various labels such as needs assessment, program monitoring, program impact, and program validation-- any of which may be included within the term _evaluation_.

The first systematic evaluation of an educational program in the United States was Joseph Rice's 1897-1898 comparative study of the spelling performance of 33,000 students in a large city school system (Worthen & Sanders, 1973, p. 2). Rice sought to revise the spelling curriculum by showing the ineffectiveness of the extended spelling drills that prevailed during that time. Since Rice, other educators have recognized the role that education can play in curriculum improvement.

It was out of a desire to improve the existing program that the Guilford County School Board decided that the reading program should be evaluated during the 1978-79 school year. A committee of

principals, teachers, and central office personnel was organized to conduct the evaluation. The committee, headed by the reading supervisor, planned, conducted, and interpreted the study to form recommendations for improving the program.

The committee decided to conduct a three-part needs assessment. For part one of the needs assessment, perceptual information was needed to describe the reading program. Perceptual information was also needed for part two of the needs assessment which was to assess the satisfaction level of principals, teachers, and parents with the reading program. For part three, test data were needed to determine how well the students were achieving in reading. Descriptive information about the reading program was collected from 722 language arts teachers and 32 principals of grades K-12 using locally developed teachers' and principals' surveys. Level of satisfaction with the reading program was obtained from a survey of the principals and teachers and from a 5% random sampling of students (N = 1,245) stratified by levels K-3, 4-6, 7-9, and 10-12. The parents of these same students supplied information about the satisfaction level of parents with respect to the reading program by responding to a locally developed Parents Survey. Archival data from the North Carolina Annual Testing Program were used to determine the impact of the reading program upon student performance.

Results from the needs assessment indicated that differences existed in basal reader series, instructional practices, and organizational patterns. Four basal reader series were being used: the Rand McNally Young American Series, the Ginn 360 series, the Holt Basic

Reading Series, and the Houghton Mifflin Series. Approximately 90% of the schools responded that the teachers depended solely on the basal reading series for reading objectives, for content, and for student evaluation. Ten percent responded that they used reading management systems based on specific objectives and criterion-referenced tests. Some of the schools reported that reading was taught within a language arts block while others reported that reading was taught at a special time outside the language arts block.

The diversity of approaches to reading instruction and the lack of continuity from grade to grade did not have an adverse effect on the students' reading achievement. The mean total reading score of third-graders was equal to the national norm, and the mean total reading score of sixth-graders was two months (+.2) above the national norm, as measured by the California Achievement Test (CAT) during the spring of 1979.

During the 1979-80 school year, the Reading Curriculum Committee met on a regular basis to plan a reading program that would provide more continuity and consistency within each school and among the schools in the school system. Each school was represented on the committee and there was a constant flow of information from the committee to school staffs with accompanying feedback from school staffs to the committee. The committee wanted to plan a program with components such as a philosophy and K-12 reading skills continuum that would reflect the thinking of the entire staff of the school system. This required planning and writing the program at the committee level, submitting the written plans to school staffs for feedback, and using the feedback to arrive at a final plan for the program.

The program was ready for implementation during the 1980-81 school year. READ was the acronym for the motto of the program: <u>R</u>eading <u>E</u>xcellence <u>A</u>ccelerates <u>D</u>evelopment. The READ curriculum consisted of:

1. a curricular guide with a philosophy for the program and objectives for grades K-12,

2. adoption of the Harcourt Brace Jovanovich Bookmark Reading Series as the lead basal reading series,

3. adoption of the Scott-Foresman <u>Basics in Reading</u> as the co-basal reading series. This series was intended for use with those students who were reading at least two reading levels below their grade level and who did not succeed with the lead basal reading series,

4. coordination of the system's reading objectives with the basal readers, the <u>Prescriptive Reading Inventory</u> (PRI) and the CAT,

5. a skills checklist to relay information about student progress in reading from grade to grade and from school to school,

6. a system for reporting growth in reading to parents,

7. a procedure for developing a centralized resource file with materials keyed to the objectives, and

8. the use of periodic and cumulative tests to determine mastery of the reading objectives.

Once the READ curriculum was implemented, objective evidence was needed to determine what effect, if any, the intervention was producing within the school system.

Statement of Purpose

The purpose of this study was to determine the effect of the READ curriculum. In order to achieve this purpose, two central issues needed to be addressed. The first issue dealt with the problem of determining READ program effects. This issue was addressed through the use of a pretest-posttest design. However, the pretest-posttest design was a single-group design with the problem that any observed effects could not be directly attributed to READ. Therefore, the second issue dealt with validation, that is, determining the accuracy of the inferences that the identified effects were due to the READ curriculum and not to other factors. A metaevaluation technique was developed and implemented as a means of dealing with this issue. Metaevaluation refers to the evaluation of evaluations or the evaluation of evaluators. The purpose of a metaevaluation is to improve the technical quality of an evaluation. In this instance, metaevaluation referred to the use of multiple independent replications to validate the causal inference that program effects were due to the READ curriculum.

Hypotheses

Four major hypotheses were tested as a part of the study. The first three hypotheses dealt with the issue of determining READ program effects while the fourth hypothesis dealt with the issue of validating that the effects were due to the READ curriculum.

It was anticipated that principals and teachers would describe the reading program differently in 1982-83 than in 1978-79; that the satisfaction of principals, teachers, parents, and students would be

greater; and that the reading achievement of the students would be
higher as a result of the READ curriculum. The following hypotheses
were tested in order to determine the effects of the READ curriculum:

Hypothesis 1. Principals and teachers will report a
greater use of practices and procedures related to the
READ curriculum in 1983 than they did in 1979. This
hypothesis will be tested at the .05 level.

Hypothesis 2. Principals, teachers, parents and
students will report greater satisfaction with the
reading program in 1983 than they did in 1979. This
hypothesis will be tested at the .05 level.

Hypothesis 3. Students in grades 3 and 6 will,
on the average, attain higher reading achievement
scores on the CAT in 1982 than students in grades 3
and 6 attained in 1979. This hypothesis will be
tested at the .05 level.

Hypothesis 4. Independent groups of school
staff will agree on changes in the reading program
that can be attributed to the READ curriculum. This
hypothesis will be tested at the .05 level.

A one-group pretest-posttest study was conducted in order to test
the first three hypotheses. The four hypotheses and design of the study
were as follows:

	1979	1981	1983
Hypothesis 1 Description of the reading			
program.	0	X	0

		1979	1981	1983

Hypothesis 2 Satisfaction of teachers,
            principals, parents and
            students with the reading
            program.   0   X   0

Hypothesis 3 Reading achievement of third
            and sixth grade students.   0   X   0 (1982)

Hypothesis 4 Validation of the causal inference
            that program changes were due to
            READ   0

In the above design, 0 = observations (data collected in 1979 and again in 1983 from principals, teachers, parents, and students) and X = intervention (implementation of the READ curriculum in 1981). This design is similar to Campbell and Stanley's (1963) pre-experimental design number two. This design lends itself more readily to local program evaluations than do other research studies for a number of reasons:

(1) Programs take on an individual nature with local qualities or idiosyncrasies relative to the local program. It may be more important to uncover local program impact than to use a more generalizable design (Kennedy, 1978). Kennedy also concluded that although the one-group study may be judgmental in nature, it falls under the general rubric of a decision-making model of educational evaluation.

(2) Stake (1976) noted that the one-group format comes close to simulating the reality of life in school settings. This is particularly true since local curriculum specialists, supervisors, and administrators are more often in the business of evaluating and improving a

total program than they are in conducting research. By the same token, achievement in reading is so important to total achievement growth that it may be unethical to withhold an improved reading program from subgroups of the client population as would be necessary in a more controlled study.

   (3) Continuous assessment and evaluation should be a part of the total curriculum system (Beauchamp, 1975). The pretest-posttest design lends itself to the collection of baseline data about the program at regular intervals.

   On the other hand, Campbell and Stanley (1963) listed several factors that can jeopardize the internal and external validity of the one-group pretest-posttest study. These factors were history (events occurring between the pretest and posttest in addition to changes in the reading curriculum that may have affected the outcome of the study), maturation of the students, statistical regression toward the mean on the posttest, and errors in instrumentation and testing. Because the design of the study did not provide controls for these factors, a metaevaluation technique was used to determine if program changes could be attributed to the READ curriculum and not to these other factors.

   The term "metaevaluation" was introduced by Scriven (1976) to refer to the evaluation of evaluations or the evaluation of evaluators. The inherent purpose of a metaevaluation, as suggested by Scriven, was to provide a quality mechanism in order to improve the theory and practice of evaluation. Cook and Gruder (1978) listed seven models of metaevaluation:

    (1) essay review of an evaluation report,
    (2) review of the literature about a specific program,
    (3) empirical reevaluation of an evaluation or program,
    (4) empirical reevaluation of multiple data sets about
         the same program,
    (5) consultant metaevaluation,
    (6) simultaneous secondary analysis of raw data, and
    (7) multiple independent replications. (p. 481)

    According to Cook and Gruder, the validity and credibility of an

evaluation is enhanced by agreement in the findings between a primary

and a secondary analysis of data and by agreement among the findings

from multiple simultaneous replications. Although Cook and Gruder

referred to evaluations conducted by outside contractors, it may be

inferred that validity and credibility would likewise be enhanced in

studies conducted by local evaluators who utilize these same techniques.

Using this same reasoning, allowing local experts to give independent

judgments about the reading program was viewed as a means of adding

validity to this study since the one-group pretest-posttest design did

not rule out history, maturation, regression, instrumentation, and

testing as threats to internal validity. Therefore, model number seven

of the Cook and Gruder metaevaluation classification, multiple indepen-

dent replications, was used to validate the causal inference that changes

in the reading program were due to the READ curriculum. Six schools

were randomly selected to replicate this part of the study. The

Reading Curriculum Committee made the seventh replication. The seven

replications were considered sufficient to get a cross section of the

various perceptions of teachers and principals about how much of READ

was implemented although the basic components of READ were available to

each school. The components were the READ Curricular Guide, similar

basal and co-basal reading materials, the Checklist of Reading Skills, and the Reporting Form for Parents.

The seven different groups of teachers were asked to derive lists of practices and procedures that were highly related to the READ curriculum and which were used more frequently since the implementation of READ. These lists were derived by having two rounds of impact questionnaires. On the first questionnaire, teachers rated practices and procedures in reading on the basis of whether they were practiced more frequently since implementing the READ curriculum. Those items that were indicated as being practiced more frequently were placed on the second questionnaire. The teachers were then asked to select the ten practices and procedures that were most closely related to the READ curriculum. Those items that were ranked as being most closely related to the READ curriculum were listed as changes in the reading program attributable to the READ curriculum. The seven independently derived lists were compared for similarity by using Kendall's Coefficient of Concordance. This part of the study was used to test hypothesis 4. However, items that met the criteria of appearing on all seven lists were considered validated effects of the READ curriculum.

Definition of Terms

The following definitions are presented in order to clarify the meaning of various terms as they are used in this study:

1. Evaluation. This term is defined in several different ways depending upon the model of evaluation used. Here evaluation is defined as the process of obtaining information and making judgments about the inputs and impact of a program

as an aid to decision making for program improvement.
This definition of evaluation is a combination of the
definitions that have been given by Stake (1967),
Stufflebeam et al. (1971), Cronbach (1973), and Guba (1972).

2. Impact statements. Outcomes or changes in characteristics,
practices, or procedures that were attributed to the READ
curriculum by the seven groups of teachers working
independently were called impact statements.

3. Metaevaluation. Metaevaluation is defined as the process
of delineating, obtaining and using descriptive and
judgmental information about the technical adequacy,
utility, ethics, and practicality of an evaluation in
order to guide the evaluation and publicly report its
strengths and weaknesses (Stufflebeam, 1981, p. 146).

4. Program. A program includes instructional materials,
procedures or management plans complete with rules for
operation and implementation (Altschuld & Hines, 1982,
p. 333).

5. Program effects. Outcomes such as greater satisfaction
and higher reading achievement that were the results of
the READ curriculum were called program effects.

6. Reading. For the purpose of this study, reading was
defined as the recognition and comprehension of written
language. (Reading Curriculum Committee of the Guilford
County School System, 1979, p. 2).

Assumptions and Limitations

Educators today recognize the necessity of program evaluation. It is one aspect of a curriculum system, including curriculum development, curriculum implementation, and curriculum evaluation (Beauchamp, 1975). The first assumption of this study was that determining program effectiveness is an integral part of a total curriculum improvement process. Whether the evaluation takes place before or after program changes, recommendations about the program are a part of the evaluation.

A part of the evaluation process includes identifying potential audiences for the evaluation report. The second assumption of this study was that the principals, teachers, parents, and students involved in the reading program were among the most important potential audience for the evaluation report. Involving these participants in the evaluation process was viewed as one way of simulating interest in the results of the study since program evaluation is of little value unless some use is made of the results. Just as important is the involvement of staff and teachers in the development and conduct of the metaevaluation. This involvement should contribute to the use of evaluation results.

The third assumption of this study was that local program evaluations which are conducted by an inside evaluator can be improved through the use of metaevaluation techniques. The purpose of metaevaluation is to improve the technical quality of an evaluation. If this is true in studies conducted by contract evaluators who ordinarily bring objectivity and expert knowledge to a study, it is likewise true in local evaluation studies that lack objectivity or control for some of the threats to validity.

The study was confined to the reading program of the Guilford County School System. The results of the study are limited in generalization to that specific population. However, to the extent that the Guilford County School System is representative of other county school systems with enrollments of approximately 25,000, the results may have meaning for such systems.

Significance of the Study

The theory and practice of metaevaluation is relatively new and there is a scarcity of literature in this area. This is particularly true of the use of group judgment methods, such as the one proposed in the present study. The adaptation and development of alternative types of metaevaluation techniques at the local school district level are concepts that have the potential for helping evaluators improve the internal and external validity of their studies without bringing in an outside evaluator, who may have no vested interest in the program; who may, because of other commitments, operate on a time frame that is not convenient for the local program participants; and most importantly, who may require more money and take more time to understand the program and the program participants. On the other hand, multiple independent replications at the local level have the potential for involving more local people in each replication. These replications can provide judgments of worth and value during the planning, implementation, and interpretation stages of curriculum improvement and curriculum evaluation.

It is hoped that this study will stimulate more interest and research into the use of metaevaluation techniques for improving local

evaluation studies as well as adding to the body of literature in the area of metaevaluation.

CHAPTER II

REVIEW OF SELECTED RELATED LITERATURE

Because there are different reasons for and different expectations from an evaluation, there are alternative conceptual frameworks outlining various types, functions and procedures of evaluation. The oldest and perhaps best known evaluation model was proposed by Ralph Tyler (1949). Since Tyler, well-known models have been proposed by Robert Stake (1967), Michael Scriven (1967), Marvin Alkin (1972), Malcolm Provus (1973), and Daniel Stufflebeam (1971) among others.

This review of the literature examined the major models or approaches to evaluation in an attempt to identify procedures that would be suitable for evaluating the effects of curricular intervention at the local school district level. In addition to examining these various approaches, the literature related to the evaluation of reading programs was reviewed and presented as each model of evaluation was described. There was also an assessment of which of the models dominated current reading program evaluations. The purposes of these reading program evaluations, as well as the kind of information gathered from them, provided additional information to guide this study.

The concept of metaevaluation was introduced by Scriven (1976) as a means of improving the technical quality of an evaluation. Since the one-group pretest-posttest design proposed for the study lacked controls for various threats to internal and external validity, the literature

related to metaevaluation was reviewed in an effort to find a way to support the causal inference that improvements in the reading program occurred because of the curriculum intervention and not because of other factors.

Thus, the complete review of selected literature has two major sections and a summary.

Major Evaluation Models

The literature related to educational evaluation is replete with various models and approaches which focus on selected features, unique functions, and various procedural patterns that relate to the purposes of evaluation. Several writers have grouped these various models and approaches according to the definition or the methodology used for the evaluation. Worthen and Sanders (1973), and House (1978) showed at least eight different classifications while Borich and Jemelka (1981) showed five. Guba (1972) in the area of evaluation and Rogers (1983) in curriculum have suggested that all of these models and approaches could be subsumed under three major definitions or models that dominate the field of curriculum evaluation. These models, based on their primary concern, may be classified as the achievement-of-desired-outcomes models, the assessment-of-merit models, and the decision-making models.

The achievement-of-desired-outcomes model is used primarily to evaluate the achievement level of individual students or groups of students. The curriculum evaluator employing this model is interested in the extent to which students are performing in accord with expected behavior.

The assessment-of-merit model is primarily concerned with the examination of merit of a given entity. This model can also concern itself with stages in the curriculum process when certain evaluative questions are raised. The stages refer to functions studied at both the formative and summative periods of the implementation of a curriculum program.

The decision-making model of curriculum evaluation is primarily concerned with future actions based on the evaluation results. This model seeks to sort out alternatives to assist in decision-making.

Differences in these three models imply different evaluation activities although the models are not mutually exclusive. Combining components of these various models may create evaluations that are more suitable for specific problem areas, according to Rogers (1983).

The purpose of the present study was to determine the effects of curriculum intervention at the local school district level. In the review which follows, each of the three major evaluation models was examined with respect to this purpose and in the light of the two major issues addressed by the study:

Determining the effects of curriculum intervention at the local school district level.

Validating that the effects were due to the curriculum intervention and not to other factors.

Achievement-of-Desired-Outcomes Model

The Tyler model. The oldest and perhaps best known model was the achievement-of-desired-outcomes model. Ralph Tyler (1949) developed this approach to evaluation during the forties and it has since become a

classic. Tyler defined evaluation in the following manner:

> The process of evaluation is essentially the process
> of determining to what extent the educational objectives
> are actually realized by the program of curriculum and
> instruction. However, since educational objectives are
> essentially changes in human beings, that is the
> objectives aim to produce certain desirable changes in
> the behavior patterns of the student, then evaluation
> is the process for determining the degree to which these
> changes in behavior are actually taking place. (p. 105)

The procedural design for the Tyler model was summarized in the

following steps:

1. Formulate educational objectives and classify them
   according to level of specificity.
2. Define each objective in terms of student behavior.
3. Identify situations in which students can be expected
   to display these types of behavior.
4. Develop or select techniques for appraising student
   behavior.
5. Gather and interpret performance data (Tyler, 1949,
   p. 499).

The Tyler model demonstrated that the central concern of evaluation

was student performance. No provisions were made for examining the

design or implementation of the program of instruction, nor did this

approach judge the worth or the value of the program. The question of

why objectives were or were not achieved was likewise not addressed by

this model.

Examples. The evaluation of the reading program of Reading,

Massachusetts, was an example of this kind of evaluation (Mason, 1981).

The evaluation was conducted by the Director of Reading in conjunction

with the reading teachers and the building principals. Approximately

ten percent of the students in each grade (2-12) participated in the

evaluation. Five goal statements were the criteria for evaluating the

program. Test scores and data from a student questionnaire were used

to draw conclusions about the program. Mason (1981) reported that the students were developing a high degree of proficiency in the various aspects of reading: word recognition skills, comprehension, reading skills needed to function in society, reading skills for individual purposes, and students' attitudes toward reading.

In a similar manner, the evaluation of the Division of Special Education and Pupil Personnel Services (DSEPPS) 1975-76 Supplementary Reading Program for Handicapped Children reported findings in terms of student mastery of objectives. The program served 1,578 children (5-16 years old) in 43 schools in New York City. Ramsey (1976) concluded that the program produced statistically significant improvement in children's reading levels, that for the majority of children, partici- pation in the program resulted in the mastery of instructional objectives which were failed on the pretest, but that it was not demonstrated that 70% of any of the groups of children were able to master eight or more instructional objectives from the California Prescriptive Reading Inventory. Tables with statistical data were provided for illustrating and validating the program effects.

Both the Mason and the Ramsey studies were examples of evaluation reports where the achievement of desired outcomes in terms of students' growth in reading was examined. Findings were reported in terms of reading grade level and mastery of instructional objectives. The studies limited their findings to one aspect of the reading program, the effect of the program on student achievement. While this is a paramount concern of educators, the improvement of instructional practices, reading materials, teachers' knowledge and other factors are all

important aspects of a reading program. Other techniques would have
to be used in addition to the achievement-of-outcomes model in order
to take a look at these aspects of the program.

Huebner (1967) recommended that program evaluation in reading
should ideally encompass the total range of the reading program
including materials and techniques of instruction. Heubner suggested
the use of the Reading Survey as a means of acquiring this complete
perspective of a current reading program. She further suggested that
the survey be formulated, conducted, and interpreted locally. The
results of the Reading Survey could well lead to suggestions for
changes or could confirm that the reading program may already be the
best possible program.

Summary. Use of the Tyler model for determining the effects of
curriculum intervention would not accommodate the full range of the
reading program as described by Heubner but would limit the findings
to student performance.

The Assessment-of-Merit Models

Worthen and Sanders defined evaluation as ". . . the determination
of the worth of a thing" (1973, p. 19). In a similar manner,
Michael Scriven defined evaluation as "a methodological activity that
consists in the gathering and combining of performance data with a
weighted set of criteria scales to yield either comparative or numerical
rating." (1972, p. 123).

The Scriven model. Scriven distinguished between the goal of
evaluation (to judge the merit of something) and the role of evaluation
(constructive use of evaluative data). In his earlier work, Scriven

noted two main evaluative roles: formative, to assist in developing
curricula; and summative, to assess the merit of curricula once they
have been developed. He also distinguished between intrinsic and
payoff evaluation. Intrinsic evaluation appraised the qualities of
a teaching instrument regardless of its effects, by assessing such
factors as content, goals, grading procedures, materials and teacher
attitude. Payoff evaluation was concerned not with the nature of the
teaching instrument but, rather, with its effects on students. Both of
these could serve either formative or summative roles.

Scriven introduced and described the concept of goal-free evalua-
tion, in which the evaluator remained ignorant of a program's written
goals and searched for all effects of a program regardless of any
rhetoric concerning what the program was intended to produce. The
advantage of goal-free evaluation was that important unanticipated
effects might be discerned that the goal-based evaluator would miss
because of a preoccupation with stated goals. Goal-free evaluation
would be more objective since the goal-free evaluators did not allow the
program staff to orient them concerning the program's intent. Scriven
presented goal-free evaluation as a supplement to goal-based evaluation.

Scriven (1974) extended his earlier ideas and subsequently
formulated the Pathway comparison model. The rationale for this model
was that evaluation essentially is a data reduction process that obtains
and assesses large amounts of data and then synthesizes them all into
an overall judgment of merit. The nine steps included the following:

1. characterizing the nature of the program to be
   evaluated;
2. clarifying the nature of the conclusion wanted from
   the evaluation;

3. assessing evidence about cause and effect relationships between independent and dependent variables in the program;
4. comprehensively checking for all consequences of the program;
5. determining and assessing the criteria of merit and the philosophical arguments pertaining to the programs;
6. assessing various kinds of program costs;
7. identifying and assessing the program's critical competitors;
8. identifying the program's constituents and performing a needs assessment to determine the program's potential impact; and
9. forming a conclusion about the merit of the program (Scriven, 1974, pp. 101-102).

These steps were not intended to be performed in any particular sequence but all had to be completed before the Pathway model was properly implemented. An evaluator might also cycle through the model several times during the evaluation of a program. Early cycles were formative and the final cycle was summative in Scriven's terminology.

The Stake model. Robert E. Stake (1967) defined evaluation as the process of fully describing and fully judging the merits of an education program. The design of the Stake model can be examined by looking at the two acts of evaluation: description and judgment. The descriptive act may be summarized in the following steps:

1. The rationale or philosophy of the program was examined.

2. The program was examined in terms of its antecedents, transactions, and outcomes. The antecedents were the conditions existing prior to teaching and learning which may relate to outcomes. These would include resources available, status of the participants, and other preprogram data. The transactions were the succession of encounters and engagements which are compromised by

the process of education. The outcomes are the
results or products of an educational experience.

3. The antecedents, transactions, and outcomes were
    examined for the contingencies between them. By
    contingency Stake meant the logical and empirical
    connection between them. These contingencies needed
    to be examined at two levels: first, what was
    intended, and, secondly, what was observed. The
    question of contingency sought to discover the
    connecting relationships between intended transactions
    and intended outcomes, as well as between observed
    transactions and observed outcomes.

4. The evaluator examined the congruency between what
    was intended and what was observed in the three
    areas of antecedents, transactions, and outcomes.
    The relationship between intents and observations
    was congruent when what was intended actually
    happened. If what was intended did not happen,
    then the relationship was incongruent.

The second act of evaluation was judgment. Under the Stake model,
judgment was reached in the following manner:

1. Standards were generated. Standards of excellence were
    needed in the three areas of antecendents, transactions,
    and outcomes for a comprehensive evaluation of the
    educational program.

2. The standards were compared to the observations and intents to determine whether standards had been met.

3. Judgments were made upon these standards. Based on the standards which were given importance, a program or an aspect of a program was judged whether or not it was worthwhile.

Examples. Studies that made use of the assessment-of-merit model were reported by Wallen and Wisely (1970), Mills and Crawford (1973), and Ellis (1975).

Wallen and Wisely (1970) reported that the Winston-Dillard School District in Eugene, Oregon, used a research team from the University of Oregon to conduct a context evaluation and came up with recommendations for improving the reading program of their school system. Data were obtained from a survey of 47 teachers, from reading achievement tests, and from observations of reading materials.

The Mills and Crawford (1973) report on the evaluation of the reading-language arts program of the Fort Gay-Thompson Ohio Schools appeared to assess the merit of the program within the framework set forth by Stake. The evaluation report contained data drawn from two major sources: intent--the planned-for conditions, behaviors, and effects; and observations--surroundings or events observed in a direct or personal way by the evaluator. Two categories of information were investigated within each source: antecedents or conditions existing prior to teaching and learning, and transactions meaning encounters of students with teachers or students; the succession of engagements which the educational process comprises. Specific antecendent conditions

viewed as most critical to the program's effectiveness were student, teacher, and parent characteristics. Selected transactions were curricular content and context, instructional materials, physical plant, and school organization. Some of the evaluation findings were that primary grade students viewed reading primarily as form rather than function, elementary students considered reading as more important than did secondary students, teachers regarded using reading to learn something the student wanted to know as the most important objective, and parents' expectations for their children were higher than the expectations of teachers or students. Recommendations based on the evaluators' judgments were an outcome of the study.

Ellis (1975) stated that the Massachusetts Advisory Council on Education decided to evaluate ten of their most successful inner-city schools in order to determine what made them successful. The Education Research Corporation (ERC) was contracted to carry out the evaluation. An accreditation type of evaluation was carried out. A group of 25 experts   from various relevant fields such as reading, measurement, administration, and individualized instruction assisted the ERC in preparing for and conducting visits to the schools. Recommendations for action were an outcome of the study.

Many value judgments were involved in program evaluation when the assessment-of-merit model was used. Value judgments were made explicit in the selection and the definition of the problem as well as in the development and implementation of the procedures of the study.

Taylor and Maguire (1972) pointed to five groups as judges who should be heard regarding education: people for society at large,

subject-matter experts, teachers, parents, and the students themselves.
According to Taylor and Maguire, superficial polls, letters to the
editor, and other incidental judgments were insufficient. In their
opinion, an evaluation of a school program should portray the merit
and fault perceived by well-defined groups and should be systematically
gathered and processed. Thus, judgment data and description data were
seen as essential to the evaluation of educational programs.

   Summary. The assessment-of-merit models, in a number of areas,
showed a marked improvement over the achievement-of-outcomes model for
determining program effects. This model provided for the evaluation
of content, goals, grading procedures, and materials in addition to
achievement and attitudes. The issue of values and standards of
judgment were also addressed. In general, a more comprehensive look
at a program and   its effects was accomplished through the use of
this model.

Decision-Making Models

   Daniel L. Stufflebeam et al. (1971) stressed the role of evaluation for
decision-making and program change. Stufflebeam defined evaluation as
"the process of delineating, obtaining, and providing useful information
for judging decision alternatives" (p. 19). Stufflebeam's ideas are
embedded in the CIPP model of program evaluation.

   The CIPP Model. CIPP is an acronym which stands for the four types
of evaluation that were included in this model. These were Context,
Input, Process, and Product evaluations.

   1. Context evaluation defined the relevant environment,

      described the desired and actual conditions pertaining to

the environment, identified unmet needs and unused
opportunities, and diagnosed the problems that
prevented needs from being used. Context evaluation
served planning decisions to determine objectives.

2. Input evaluation provided information for determining
   how to utilize resources to meet program goals. It
   involved identifying and assessing relevant
   capabilities of the responsible agency, alternative
   strategies for achieving program goals, and alternative
   designs for implementing a selected strategy. Input
   evaluation served structuring decisions to determine
   program designs.

3. Process evaluation provided periodic feedback to
   persons responsible for implementing plans and procedures.
   The strategy involved identifying and monitoring
   continuously the potential sources of failure in a
   project, projecting and servicing preprogrammed
   decisions, and describing what actually occurs
   during the program. Process evaluation served
   implementing decisions to control project operations.

4. Product evaluation involved measuring and interpreting
   attainments, not only at the end of a program or
   project cycle but as often as necessary during the
   program.

Evaluation has generally focused on products, but in contrast to
current usage, product evaluation was understood as a continuous activity

since attainments occur throughout the program. It served recycling decisions to judge and to react to project outcomes. It served the decision maker who must decide whether to continue, to terminate, or to modify a program.

The four kinds of evaluation in the CIPP model were designed to service four types of decisions, respectively. These four types of decisions were generated by crossing two dimensions of decisions: the function of the decision, whether it pertains to ends or means and the relevance of the decision, whether it is directed towards intentions or actualities. It was assumed that all educational decisions could be classified into one of these four types. Planning decisions were concerned with intended ends and structuring decisions with intended means. Recycling decisions were concerned with actual ends and implementing decisions with actual means. Each of these four decision types was served by a different evaluation strategy, as has been stated above. Planning decisions were served by context evaluation, structuring decisions by input evaluation, implementing decisions by process evaluation, and recycling decisions were served by product evaluation.

Systems assessment. Marvin C. Alkin (1972) defined evaluation as the process of ascertaining the decision areas of concern, selecting appropriate information, collecting and analyzing information in order to report summary data useful to decision-makers in selecting among alternatives.

Alkin identified five areas in which decisions must be made about the state of the system.

1. Systems assessment was described as a means of determining
   the range and specificity of educational objectives
   appropriate for a particular situation. The needs
   represented a gap between the goal and the present
   state of affairs. The evaluative problem, then, became
   one of assessing the needs of the students, the needs
   of the community, and the needs of society in relation
   to the existing situation. Assessment, therefore, was
   a statement of the status of the system as it existed
   in comparison to desired outputs or stated needs of the
   system.

2. Program planning was concerned with providing information
   which would enable the decision-maker to make planning
   decisions, to select among alternative processes in
   order to make a judgment as to which of them should be
   introduced into the system to fill most efficiently
   the critical needs previously determined.

3. Program implementation was concerned with determining the
   extent to which the implemented program met the
   description formulated in the program planning
   decision. In the case of an existing program where
   no known changes had been implemented, the evaluation
   task at this stage was to determine the degree to which
   planning descriptions of the program coincided with the
   implemented program and the extent to which assumed

descriptions of inputs to the system (students)

corresponded with observed inputs.

4. Program improvement was an expected outcome of the

implementation process. As the evaluator identified

problems and collected and analyzed related information,

data were presented immediately to the decision-maker

so that changes could be executed within the sytem to

improve the operation of the program.

5. Program certification was needed to provide information

that would enable the decision-maker to make decisions

about the program as a whole and its potential general-

izability to other situations. The evaluator might

attempt to provide information which would enable the

decision-maker to determine whether the program should

be eliminated, modified, retained, or introduced more

widely.

The CIPP model and systems analysis represented total systems

approaches to evaluation because they served both in planning and in

implementing an educational program. They provided for evaluating

process with continuous feedback. Besides evaluating process, including

group process, these models also clearly provided for evaluating

procedures both in terms of the selection of procedures through input

evaluation and in their implementation through process evaluation. While

the question of values was not as predominant as in the assessment-of-

merit model, it was not neglected as in the Tyler model. The evaluator

was urged to examine the values served by various decision alternatives.

It was not the task of the evaluator to generate standards through empirical methods, as in the Stake model. Rather, it was the evaluator's task to provide information to the decision-maker regarding those values which would be served by alternative decisions but not to offer judgments himself.

Context evaluation. Robert L. Hammond (1971) described context evaluation as a good way of assessing the degree to which reading programs were accountable. He defined evaluation as, "the process of delineating, obtaining and providing useful information for judging decision alternatives" (p. 31). According to Hammond, context evaluation

> defines the relevant environment, describes the actual conditions pertaining to the environment, identifies needs and unused opportunities and diagnoses the problems that prevent needs from being met and opportunities from being used. (p. 31)

The Hammond framework was similar to the context evaluation described by Stufflebeam in the description of the CIPP model.

A number of state and local school districts--Oregon, Washington, and Wisconsin to name a few--have used the Right To Read Assessment Planning Handbook, which was designed by the U. S. Office of Education (1974) as a guide for planning more effective programs. This handbook included needs-assessment materials combined with a self-study guide by which a committee or task force might follow a step-by-step procedure for collecting data and making decisions.

1. Identify the population.

2. Assess the current program status.

3. Identify and prioritize new objectives.

4. Review effective programs and/or program components.

5. Plan program of diagnosis, prescription, and evaluation.

6. Identify instructional approaches, methods, and
   techniques.

7. Plan staff development.

8. Identify needed personnel, materials, services, and other
   costs.

9. Provide for continuous evaluation throughout the year.

This systematic approach to the evaluation of a reading program closely resembled the systems approach described by Alkin and contained all of the major steps outlined in the CIPP Model.

The discrepancy model. The evaluation model formulated by Malcolm Provus (1971) was generally considered to be a decision-making model although judgment played a very large role. In this model, evaluation was described as the process of agreeing upon program standards, determining whether a discrepancy exists between some aspect of the program and the standards governing that aspect of the program, and using discrepancy information to identify the weaknesses of the program. The decisions that stemmed from this model related to either improving, maintaining or terminating a program.

Examples. The Pittsburgh (Pennsylvania) Public Schools maintained an evaluation unit in the school system's department of research patterned after the discrepancy model. The evaluation model was composed of five stages: program design, program operation, interim products, terminal products, and costs. Comparisons of performance to standards were developed for each stage (Provus, 1971).

The Smyrna School District in Smyrna, Delaware, also used a discrepancy model to evaluate the reading program. McCormick (1976) reported that the existing state of the reading program was compared to "Standards of Excellence for Reading" in Delaware. Needs were identified and ranked in order of priority. A corrective plan was recommended by a district-level committee. An ERIC search of reading program evaluations showed that many other states have "Standards of Excellence" or "Criteria for Excellence in Reading," including Alaska, Pennsylvania, Tennessee, Iowa, New York, Alabama, Texas, Louisiana, Georgia, West Virginia, and North Carolina.

McGuire (1968) reported a statewide status study of reading programs in Rhode Island. Data concerning reading achievement were correlated with factors such as class size, instructional approach, and family income. Successful practices, as well as indicated needs, were an outcome of the assessment. This method was a form of systems assessment.

In a similar manner, Gaberina (1976) used a systems approach to evaluate the reading program of the Pennsbury School District in Fallsington, Pennsylvania. Gaberina reported that he used a mixture of "Right-to-Read Program Planning Procedures," Pennsylvania's "Generic Planning Process" and Stufflebeam's CIPP model. His final evaluation model contained context, input, process, product, and installation. Gaberina recognized the importance of the local administrator as a change agent who initiates, encourages, supports, observes, and evaluates programs. He suggested the use of the systems approach as a guide for local administrators to provide a systematic and objective means of evaluating an ongoing reading program.

The _Evaluation System Report_ of the Public Schools of the District of
Columbia (1979) was similar to the Rhode Island and Pennsbury studies.
The evaluation system was designed to combine test data with program
data and to combine both of these with a variety of other data measures
such as student absences, sex, age, size of class, involvement of aides,
and many other such measures. This procedure enabled the District of
Columbia Public Schools to follow particular students over a period of
years, to examine long-term effects of different programs, to measure
changes in program characteristics from year to year, to evaluate the
effectiveness of programs, and to identify characteristics of programs
that seem to increase effectiveness.

The Saginaw (Michigan) School District in conjunction with Ohio
State University installed an evaluation unit in the Saginaw Public
Schools. According to Hock, Sellers, Blatt, and Gault (1971), two
mechanisms guided the operation of the Saginaw unit. The first was a
model of evaluation involving context, input, process, and product
evaluation that provided both continuous and ad hoc evaluation
capabilities. The second was a policy handbook covering the planning,
programming, budgeting, organizing, directing, and controlling of the
work of the evaluation unit.

Benjamin (1978) described the Evaluation Improvement Program (EIP)
materials developed under the direction of the California State Depart-
ment of Education to be the most effective means of assisting school
administrators, program coordinators, and instructional staffs in
developing plans for Early Childhood Education and Title I programs.
The process included needs assessment combined with the formulation of

goals and objectives that were to be attained as well as structure, content, and processes of instruction.

Summary. The decision-making models seem to dominate the literature with respect to the evaluation of reading programs. The CIPP Model, the Systems Assessment Model, and the Provus Discrepancy Model described earlier were reported. In addition, the Right-to-Read Assessment and Planning Handbook (U.S. Office of Education, 1974), California Improvement Program, and combinations of these and other models were described. The number of states which have adopted standards of excellence and which use the Right-to-Read Handbook contributed to the large number of studies using decision-making models and hence the conclusion is drawn that decision-making models seem to dominate the literature with respect to the evaluation of reading programs.

For evaluating the effects of curriculum intervention at the local level, the product evaluation described in the CIPP Model and program certification of the Systems Assessment Model appear appropriate. Product evaluation provides the measurement and interpretation of the degree to which program objectives were attained while program certification deals with overall program effectiveness. Program effects identified within the framework of these two kinds of evaluation would be similar to those identified by the assessment-of-merits model in that they would not necessarily be limited to student achievement.

Metaevaluation

Definitions

The term metaevaluation was introduced by Scriven (1976) to refer to the evaluation of evaluations or the evaluation of evaluators. The inherent purpose of a metaevaluation as implied by Scriven was to

provide a quality control mechanism in order to improve the theory and
practice of evaluation.

Stufflebeam (1981), in a similar manner, defined metaevaluation as

the process of delineating, obtaining and using
descriptive and judgmental information about the
technical adequacy, utility, ethics, and practicality
of an evaluation in order to guide the evaluation and
publicly report its strengths and weaknesses. (p. 146)

The Scriven and Stufflebeam concepts of metaevaluation differed from
that of Cook and Gruder (1978) who referred to metaevaluation as "the
evaluation of summative evaluations--studies where the data are
collected directly from program participants within a systematic design
framework" (p. 470).

Cook and Gruder justified their definition with the assumption that
metaevaluations which are conducted during the assessment of a curriculum
can provide the most useful diagnostic feedback to decision makers.

Purposes

Smith (1981) stated that metaevaluation may focus on several aspects
of an evaluation study: its design, management, instruments, data
results, impact, personnel, setting, purpose, reporting, or any combination
of these. It may also focus on a single evaluation study or a collection
of studies in order to accomplish one or more of the following purposes:

1. to assess the quality, impact, or utilization of

    evaluation work,

2. to study the nature of the evaluation process,

3. to redress a possible evaluation abuse,

4. to certify evaluation work, providing for accountability

    in evaluation,

5. to illuminate and control for bias in evaluation work,

   or

6. to assess the utility of new approaches to evaluation

   (p. 267).

The purposes of metaevaluation enumerated by Smith are consistent with the purposes proposed by Scriven, Stufflebeam, and Cook and Gruder: to improve the technical quality of an evaluation.

Models

Cook and Gruder (1978) described seven metaevaluation models which they generated from an analogy with a three-factor analysis of variance. One factor was whether the metaevaluation took place simultaneously with the primary evaluation or after it; the second described whether the primary evaluation data were or were not manipulated by the evaluator; and the third referred to the number of independent data sets that could be used to evaluate a particular program. Cook and Gruder presented seven models:

1. essay review of an evaluation report,

2. review of the literature about a specific program,

3. empirical reevaluation of an evaluation of program,

4. empirical reevaluation of multiple data sets about the

   same program,

5. consultant metaevaluation,

6. simultaneous secondary analysis of raw data, and

7. multiple independent replications (1978, p. 481).

All of the Cook and Gruder models were utilized in the national "Follow Through" evaluation. St. Pierre (1982) described these models

of metaevaluation and related them to the corresponding Follow Through
evaluation studies. The House (1978) critique of the national
evaluation was an example of the first model. The work of Haney
(1977, 1978) was an example of the second model-- a review of the
literature about a specific program. Haney and Pennington's (1978)
work in the area of reanalyzing data collected on parent and teacher
attitudes fell into the third category, while the fourth model,
empirical reevaluation of multiple data sets, about the same program,
was exemplified by the work of Goodrich and St. Pierre (1979). The
U. S. Office of Education reviewed the evaluation and provided ongoing
feedback on the progress of the evaluation. This was viewed as an
example of the fifth model, consultant metaevaluation. Another study
by St. Pierre (1978) was an example of the sixth metaevaluation model,
simultaneous secondary analysis of raw data. Several analyses of the
data were performed which led to somewhat different conclusions from
those presented in the national report. The seventh model of
metaevaluation research, multiple independent replications, referred
to an approach in which separate groups carried out independent
evaluations. Some of the Follow Through sponsors used portions of
their funding to conduct internal evaluations that were independent
of the external national study. These were examples of multiple
independent replications. Smith (1981) listed ten possible forms of
metaevaluation:

     1. collective professional discussions

     2. secondary data analysis studies

     3. performance audits

    4. application of formal standards

    5. comparative empirical studies

    6. research on evaluation methods

    7. administrative review procedures

    8. public review hearings

    9. methodological critiques

   10. formal criticisms (pp. 267-268).

    Smith suggested that this list was neither exhaustive nor mutually exclusive and that either form may focus on any of the aforementioned aspects of an evaluation study. There is considerable overlap in the Cook and Gruder models and the models listed by Smith.

Criteria

    Stufflebeam and the Phi Delta Kappa Study Committee (1971) sought to develop an appropriate set of criteria for metaevaluation for judging evaluation designs and reports. The committee listed four scientific criteria related to the "goodness" of an evaluation: internal validity, external validity, reliability, and objectivity. With respect to the "utility" of evaluation, six practical criteria were listed: relevance, importance, scope, credibility, time lines, and pervasiveness. One prudential requirement, efficiency, was also listed.

    Scriven (1974) and Millman (1981) have also listed criteria for evaluating evaluations. Scriven suggested thirteen checkpoints: need, market, sizes and kinds of effects, causation, impacted populations, durability, generalizability, statistical significance, legality/morality/enjoyability, cost, future availability/improvements/cost, comparative significance, and overall value. Millman used the Scriven criteria

to develop a checklist for checking the quality of the assessment of the program or product by the evaluator and for critiquing the evaluation as a product. The Millman checklist contained six cells:

1. Concerning the quality of the evaluators' assessment
   (or providing information on):
   Cell I - need, market and dissemination plan
   Cell II - performance data
   Cell III - cost and benefit
2. Concerning the evaluation:
   Cell IV - meeting the preconditions of the evaluation
   Cell V - the effects of the evaluation
   Cell VI - the cost and benefit of the evaluation
   (Millman, 1981, p. 311).

Beck (1981) reported that a study committee of the International Reading Association (IRA) had adapted a set of standards for evaluations for the improvement of reading program evaluations. The standards include specifications for the purpose(s) of the evaluation, focus of the evaluation, procedures followed in conducting the evaluation, sources of information used in the evaluation, and conclusions/recommendations based on the evaluation results. When the IRA standards were compared with the standards set by the Joint Committee on Standards for Educational Evaluation (1981), they appeared to be more useful for writing the evaluation report than for improving the technical quality of the evaluation study. In contrast, the standards set by the Joint Committee include standards of accuracy, utility, propriety and feasibility in the planning, conduct, interpretation, and reporting stages of the evaluation.

Most local evaluative studies are not now, nor are they likely to be in the immediate future, readily tailored to accommodate all the standards for evaluations proposed by Stufflebeam, Millman, Beck, and

the Joint Committee on Standards for Educational Evaluation. However, metaevaluation techniques used during the planning, implementation, and interpretation stages may serve a purpose in helping local evaluators meet many of these criteria. Cook and Gruder (1978) have suggested that the validity and credibility of an evaluation is enhanced by a metaevaluation technique such as multiple simultaneous replications. Although Cook and Gruder referred to evaluations conducted by outside contractors, it may be inferred that validity and credibility would likewise be enhanced in studies conducted by local evaluators who utilize the same technique. The present study extends the use of multiple independent replications to a local evaluation study as a means of adding to the validity of the causal inference relating program effects to curriculum intervention.

The main advantage of simultaneous independent replications is that validity and credibility are maximal if the results replicate (Cook & Gruder, 1978). Because of budget restraints, replications usually require smaller studies with the additional advantages that evaluators have greater control over how the treatments are implemented, may be able to respond more quickly and flexibly to field problems, and may be in a better position to measure and describe the treatment and processes that may have mediated any observed effects.

There are also disadvantages to using multiple independent replications. The cost may be high if the replications are large. This was not a factor for the present study since the replications were kept small. A second disadvantage of multiple independent replications is the risk of reduced quality control. Keeping the replications small

may help to hold quality at a high level. Finally, simultaneous replications have the disadvantage that the evaluator cannot learn from a consecutive replication. Consecutive replications may be a useful tool for local evaluators who may need to improve the technical quality of the evaluation while it is in progress.

Summary

This review of selected literature related to evaluation, the evaluation of reading programs, and metaevaluation has contributed to an understanding of how the three major evaluation models have been used in the field of reading and how metaevaluation can be used to improve local evaluation studies. These uses may be summarized in the following points:

1. Use of the achievement-of-outcomes model for determining the effects of curriculum intervention limited the finding to achievement.

2. Use of the assessment-of-merit models for determining the effects of curriculum intervention provided for the evaluation of content, goals, grading procedures, materials, achievement, and attitudes.

3. Use of the decision-making models for determining the effects of curriculum intervention provided for findings related to the measurement and interpretation of the degree to which program objectives were attained under product evaluation or to overall program effectiveness under program certification.

4. Decision-making models of evaluation dominated the
    evaluation of reading programs.

5. For determining the effects of curriculum intervention
    at the local level, the assessment-of-merit model
    appeared more appropriate because it took a more
    comprehensive look at the reading program and beacuse
    more of the judges who should be heard regarding
    education could have input into the evaluation.

6. Metaevaluation techniques add to the validity,
    objectivity, and credibility of an evaluation.

CHAPTER III

METHODOLOGY

The purpose of this study was to determine the impact of curricular intervention on the reading program of the Guilford County School System. In order to achieve this purpose, it was necessary to address two central issues. The first issue dealt with the problem of determining the effects of the READ curriculum, while the second issue dealt with the problem of validating the causal inference that the effects were due to READ and not to other factors. The three major hypotheses that were tested in order to address the first issue were related (a) to the description of the program, (b) to the satisfaction of teachers, principals, parents, and students with the program, and (c) to the reading achievement of third and sixth grade students. In order to address the second issue, a fourth hypothesis was tested to determine whether multiple independent replications of impact questionnaires would produce the same list of program effects attributable to the READ curriculum.

This chapter will discuss the subjects, instrumentation, and procedures used for the study. The procedures used to collect the data and to test the major hypotheses are presented separately for the two major issues addressed by the study.

Subjects

The subjects for the study were students, parents, teachers
principals from the school system of Guilford County, which is located in
north central North Carolina near the center of the Piedmont plateau.
Guilford County is essentially rural and suburban
with a relatively stable population. Data obtained from the Office of
Research and Planning evaluation indicate that the racial ratio is
approximately 82% white and 18% nonwhite for the total school system.
These percentages are not consistent from school to school. The percentage
of nonwhites in the schools range from 0.96% to 39.53%.

Of the 33 elementary schools, 24 are eligible for Chapter I
Remedial Reading Programs because they have a high percentage of low
income families. The overall percentage of students in Chapter I ranges
between 10 and 12. The percentage of parents who had not completed
high school ranges from 10 to 38%. The demographic data for students in
grades three and six have remained relatively stable from 1979 to 1982
as indicated in Table 1.

The study was conducted by the Reading Supervisor along with the
Reading Curriculum Committee which was composed of one representative
from each of the school system's 44 schools. These committee members
were selected by their principals as people who were knowledgeable about
reading and who wanted to participate in the curriculum improvement
process. Several members of the committee were specialists in reading,
holding masters degrees; others were regular classroom teachers who
taught reading along with other subjects. Junior high language arts
teachers and high school English teachers were members of the committee
as well. Each member of the committee was encouraged to form reading

curriculum committees within their own schools to help with the evaluation
of their school's program as well as with the evaluation at the system
level.

Table I

Summary of Longitudinal Data About Grades Three and Six of the

Guilford County School System

Variable	Grade	Year			
		1979	1980	1981	1982
% Minority	3	19.86	17.04	16.67	17.00
	6	17.43	15.16	14.27	16.95
Mean IQ	3	102.57	102.30	104.10	101.48
	6	101.19	101.70	102.30	100.57
% Chapter I	3	15.66	12.85	15.05	16.62
	6	11.76	8.00	7.47	10.00
% Parents Not Completed High School	3	13.52	15.62	17.90	20.04
	6	20.62	17.58	14.37	22.90

The committee decided to conduct a survey of principals, teachers,
parents, and students to tap their perceptions of the reading program.
For the pretest observations, all surveys were distributed to the princi-
pals on March 15, 1979. The principals distributed the surveys to the

teachers, students, and parents. The surveys were completed and returned by March 21, 1979. The percentage of response was as follows: principals 76%, teachers 83%, students 92%, and parents 62%.

For the posttest observations, all surveys were distributed on February 2, 1983. Members of the Reading Curriculum Committee distributed the surveys. Surveys were completed and returned by March 2, 1983. The percentage of responses for 1983 was as follows: principals 100%, teachers 100%, students 86%, and parents 70%.

Table 2 presents a summary of the total number of principals, teachers, parents, and students who were sent surveys, the number who responded, and the response rate.

Table 2

Subjects and Response Rates for 1979 and 1983 Reading

Program Survey Administration

Group	Total Population		Number Surveys Sent		Number of Responses		% Response Rate	
	1979	1983	1979	1983	1979	1983	1979	1983
Principals	44	44	44	44	32	44	76	100
Teachers	772	602	712	270	641	270	83	100
Students K-3	7,101	6,331	525	450	504	417	96	93
Students 4-12	18,238	17,660	720	870	645	714	90	82
Parents*			1245	1325	754	918	62	70

*One survey was sent to the parents of each student selected to respond to the students' questionnaires.

Sufficient numbers of surveys were returned by each group to provide valid data for analysis and interpretation. In fact, the response rate was attributed to the fact that the principals and members of the Reading

Curriculum Committee were leaders of Reading Curriculum Committees in their own schools and perhaps saw the need for and encouraged active participation of the other respondents.

While inferences about satisfaction with the reading program were made on the basis of information from students in all grades of the school system, inferences about reading achievement were made on the basis of reading achievement scores from two student groups. The two student populations selected to determine reading achievement were the third and sixth grade students from twenty-one elementary and middle grade schools. Schools were selected from which data were available for both years 1979 and 1982. These grades were selected because they were routinely administered the California Achievement Test each spring as a part of the North Carolina Testing Program. Students with learning problems were tested along with their classmates and their scores were included in the system's scores. The number of students on whom achievement data were available in grades three and six in 1979 and 1982 are shown in Table 3.

Six of the 44 schools in the school system were randomly selected to participate in the metaevaluation study. The committees felt that the total school population should be tapped since the program was planned to reflect the thinking of the entire school system. The six randomly selected schools are listed in Table 4. Six teachers from each of these schools participated in the metaevaluation.

Table 3

Number of Students on Whom Reading Achievement Data Were

Available in Grades 3 and 6 in 1979 and 1982

School	GRADE			
	3		6	
	1979	1982	1979	1982
1	99	110	163	154
2	91	93	106	109
3	63	67	98	72
4	67	61	76	68
5	58	58	68	81
6	50	39	123	107
7	47	71	175	165
8	187	192	123	107
9	73	71	175	175
10	43	41	47	49
11	53	69	102	63
12	51	55	83	78
13	21	53	123	60
14	108	18	76	31
15	69	72	67	113
16	76	60	26	51
17	61	70	115	81
18	110	117	63	153
19	33	28	44	34
20	110	100	127	126
21	57	42	161	125
	1527	1487	2135	1712

Table 4

Grade Levels and Number of Teachers for Schools

Participating in the Metaevaluation

School ID Number	Grade Taught	Number of Teachers	Teachers Participating
320	K-6	18	6
334	K-3	23	6
336	K-5	27	6
360	K-5	19	6
368	K-6	17	6
411	10-12	53	6

Instrumentation

Surveys were developed to gather perceptual data related to the description of the reading program and the satisfaction level of principals, teacher, students, and parents with the program. These surveys were developed for the initial administration in 1979. The Principal and Teacher Surveys were designed to provide information about the description of the program including questions related to organizing for instruction, goals and objectives, planning and instruction, instructing the students, methods of teaching, and evaluation practices.

The surveys for parents, students K-3, students 4-12, principals, and teachers all contained questions regarding satisfaction with the reading program. Copies of the reading surveys can be found in Appendices A-E.

The instrument used to assess reading achievement was the California Achievement Test (CAT) (CAT, 1977) (North Carolina annual testing program format). Form C, level 13 was administered to the third grade and form C, level 16 was administered to the sixth grade in 1979 and in 1982. The CAT is a well accepted, nationally normed and validated set of achievement tests with high reliability and validity ratings. The CAT was reported as having alternative-form reliability coefficients for reading ranging from .80 to .91 and was therefore judged to be suitable for group measurement. Selected by the North Carolina State Board of Education in 1978 as the basis of the State Annual Testing Program, the CAT was considered the battery most nearly parallel to the instructional program in the state. The Reading Curriculum Committee has matched the READ curriculum objectives and CAT items for all levels of the CAT. The CAT was therefore deemed valid to measure the reading achievement of the third and sixth grade student population.

<div align="center">Procedures</div>

The procedures are presented in two sections. The first section is concerned with determining the effects of the READ curriculum and the second section is concerned with validating that the effects were due to READ and not to other factors.

Issue I - Determining Program Effects

The study was conducted on the basis of perceptual and archival data using a preobservation/postobservation design. One procedure was to obtain perceptual data from a variety of populations who were knowledgeable about reading instruction in the Guilford County School System before and after the planned intervention. The second procedure was to compare the reading performance of the students before and after the intervention. The third task, related to hypothesis 4, was concerned with the validation of the causal inference that any observed effects were due to implementing the READ curriculum. These tasks can be described in terms of the four hypotheses which were proposed:

		1979	1981	1983
Hypothesis 1	Description	0	X	0
Hypothesis 2	Satisfaction	0	X	0
Hypothesis 3	Reading Achievement	0	X	0 (1982)
Hypothesis 4	Validating the Causal			
	Inference			0

In this design 0 = observations (data collected from principals, teachers, parents and students): X = the implementation of the READ curriculum. The design was discussed more fully in Chapter I. In terms of evaluation models, the design above did not fit readily into any one category. Hypothesis 1 above attempted to assess practices and procedures that describe the reading program; hypothesis 2 attempted to look at the impact of the program on principal, teacher, pupil, and parent attitudes; hypothesis 3 attempted to determine desired outcomes in terms of student achievement; and hypothesis 4 attempted to validate

that effects were due to the READ program and not to other factors.
This combining of components of various models to meet specific
problem needs was recommended by Rogers (1983) and others. Rogers
also suggested that the various evaluation models were not mutually
exclusive. This was especially true with respect to summative
evaluations which may be called product evaluations as defined in the
CIPP model and program certification in the systems assessment model.

Hypothesis 1. Description of the reading program. The
description of practices and procedures related to the reading program
was obtained from Principal and Teacher Surveys in Reading designed
by the Reading Curriculum Committee in 1979. The use of reading surveys
was recommended by Heubner (1967). As Heubner suggested, reading
surveys were formulated, conducted, and interpreted locally, Those
items on the Principal and Teacher Surveys that were related to
describing the reading program are listed in Table 5. The complete
Principal Survey in Reading along with percentage of responses for 1979
and 1983 can be found in Appendix A. Similar information related to
the Teacher Survey in Reading can be found in Appendix B.

All surveys were distributed by school principals in 1979 and
by members of the Reading Curriculum Committee in 1983. Principal
Surveys were distributed to each of the principals (N = 44) in the
school system. Teacher Surveys were sent to a sample of teachers at
each school including teachers involved in reading instruction in
grades K-6, and language arts and English teachers in grades 7 to 12.
For both the pretest and posttest, the returned surveys were processed
by the statistical center at the University of North Carolina at

Greensboro using SPSS (1979, 1982) programs to obtain the frequencies
and percentages presented in this report.

Table 5

Items from Principal, Teacher, Parent, Student K to 3

and Student 4 to 12 Surveys that Related to the

Description of the Reading Program

Survey	Items
Principal and Teacher	1. Program Organization

      A. In Organizing Students for Reading
          Instruction . . . I

          1. Divide students into more than
              one group on the basis of
              reading ability (i.e., ability
              grouping).

          2. Group students on the basis of
              specific reading needs (i.e.,
              special needs grouping).

          3. Base classroom reading instruc-
              tion on the idea of whole-class
              grouping (i.e., one group).

          4. Instruct students individually
              in reading rather than in groups
              (i.e., individualized reading
              instruction).

          5. Group students for reading
              instruction on the basis of
              commonly shared interests (i.e.,
              interest grouping).

          6. Use the results of standardized
              reading achievement tests as
              measured by the State Annual
              Testing Program.

Table 5 (continued)

Survey	Items
	7. Move students from one group to another (i.e., flexible grouping) as needs vary.
	8. Group students on the basis of the group they were in last year.
	9. Organize my reading program with the help of reading specialist.
	10. Use parents, paraprofessionals and/or community resources.
	11. Organize my reading program with the help of colleagues.
	B. Concerning Goals and Objectives . . .I
	12. Use the school system's curriculum guide in reading as the source of reading program goals.
	13. List general goals for the reading program based on the assessment of the students' reading strengths and reading needs.
	14. Maintain a record keeping system to keep track of individual progress toward specific objectives.
	15. Use the basal reader as the source of reading program goals.
	C. In Planning for Skills Instruction . . .I
	16. Organize classroom reading instruction on the basis of skill levels represented in the class.

Table 5 (continued)

Survey	Items
	17. Establish a sequence of reading skills based on assessment of student reading needs.
	18. Use the basal reading series to determine the sequence in which reading skills are taught.
	19. Select practice activities that match instructional objectives.
	20. Organize the classroom with learning centers.
	D. When Instructing Students in Reading . . . I
	21. Teach reading through non-basal materials.
	22. Assign workbook pages as practice activities which match instructional objectives.
	23. Use a file of workbook pages and exercises classified by skill and level.
	24. Use workbooks as the major guide to introducing reading skills.
	25. Have teacher-pupil planning sessions characterized by considerable give-and-take.
	26. Prepare a directory of commercial reading materials available within my school.
	27. Use the State's Guide for Evaluation of Materials.
	28. Code reading materials to reading objectives.

Table 5 (continued)

Survey	Items

29. Use audio-visual materials.

II. Instructional Practices

   A. Methods of Teaching . . . I

      30. Use the basal approach.

      31. Use the language experience approach.

      32. Use the phonics approach.

      33. Use programmed instruction.

      34. Use an individualized approach.

      35. Use management systems. EXAMPLE: PRI.

      36. Use an eclectic approach.

      37. Other (list on number 73).

      38. Encourage a child to select topics he/she or a group may wish to read about.

      39. Release a child from group work to do individual reading.

III. Teacher's Evaluation Practices . . . I

      40. Use records of independent reading done by each student

      41. Consider a child's ability to discuss what he/she has heard others read aloud.

      42. Consider work in reading other than the basal program material.

Table 5 (continued)

Survey	Items

---

43. Compare achievement to behavioral objectives by means of a criterion-referenced test (the objectives state the conditions under which a child should do something which teachers can observe, to a degree which the teachers have specified).

44. Use results of an informal reading inventory.

45. Consider to what degree a student's textbook or required reading are matched to his reading level.

46. The child and his/her teacher make a "performance contract" and the teacher assesses the child's progress in completing this contract.

IV. Teacher's Knowledge

Use the following choices for questions 47-63.

A. Much, B. Some, C. Little, D. None

47. I have training in teaching reading.

48. I have knowledge of the diagnostic-prescriptive approach to reading.

49. I use a diagnostic-prescriptive approach to teaching reading.

50. I have knowledge of motivation techniques.

51. I use these techniques.

Table 5 (continued)

Survey	Items
	52. I have knowledge of our school system's curriculum guide.
	53. I use this guide.
	54. I have knowledge of management system techniques.
	55. I use a management system.
	56. I have knowledge of standardized norm-referenced tests.
	57. I use the results of standardized norm-referenced tests.
	58. I have knowledge of criterion-referenced tests.
	59. I use the results of criterion-referenced tests.
	60. I have knowledge of criteria for selecting materials.
	61. I use criteria for selecting materials.
	62. I have knowledge of the sources for reading materials.
	63. I have knowledge of the issue of accountability and its implications.

Hypothesis 1 was tested by using sign tests (Daniel, 1978). The sign
test makes use of pluses and minuses rather than quantitative metrics.
Percentage of responses to the items on the surveys that related to
describing the reading program in 1983 were compared to percentage
of responses made to the same item in 1979. Those items were assigned

a plus if the percentage of responses indicated that the item was practiced more in 1983 or a minus if the item was practiced less than was indicated in 1979. Signs were assigned consistently even if the shift in response was no more than one percent. Responses that indicated that the item was practiced occasionally or often were combined for the comparisons. No signs were assigned to ties. Further interpretation of the data was needed to determine if more responses or fewer responses on any item was an effect of READ. Although the overall hypothesis was related to describing the reading program, the sign test was used to determine significant differences in clusters of items related to describing program organization, goals and objectives, planning for instruction, methods of teaching, evaluation practices, and knowledge and practices. For example, items 1 to 11 on the Principal   Survey in Reading were clustered in order to determine whether principals perceived that there were changes in program organization (See Table 6). The probability of obtaining the five pluses or four minuses with eleven items was .5001. Since hypothesis 1 was directional, it was subjected to a one-tailed test. At the .05 level, .5001 was not significant; therefore, hypothesis 1 was rejected with respect to changes in program organization as perceived by the principals of the Guilford County School System. The other clusters of items on the Principal   and Teacher   Surveys in Reading were analyzed in a similar manner. Therefore, the sign test was applied 14 times for smaller hypotheses related to the overall hypothesis for describing the reading program. For these analyses, probabilities equal to or less than .05 (for the obtained number of minuses) were sufficient to retain the directional hypotheses.

Table 6

Example of the Sign Test Applied to the Principal

Survey in Reading

DIRECTIONS: Choose the appropriate answer for the following statements
and mark the correct box.

A. I have done this often.
B. I have done this occasionally.
C. I have not done this.
D. I have not done this but think I should.
E. Not applicable.

1983 1979 = +
1979 1983 = −
No signs are attached
to ties
*Significant at the .05
level

	Percentage of Responses										Direction Of Difference
	1979 n=32					1983 n=44					
I. Program Organization	A	B	C	D	E	A	B	C	D	E	
A. In Organizing Students for Reading Instruction . . . My Teachers											
1. Divide students into more than one group on the basis of reading ability (i.e., ability grouping).	78	13	3		3	85	10	3	3		+
2. Group students on the basis of specific reading needs (i.e., special needs grouping).	78	19			3	51	41	8			−
3. Base classroom reading instruction on the idea of whole-class grouping (i.e., one group).	9	19	59	3	9	16	27	47	5	5	+
4. Instruct students individually in reading rather than in groups (i.e., individualized reading instruction).	28	43	13			6	27	54	9	9	0

Table 6 (continued)

Item	Percentage of Responses 1979 n=32					1983 n=44					Direction Of Difference
	A	B	C	D	E	A	B	C	D	E	
5. Group students for reading instruction on the basis of commonly-shared interests (i.e., interest grouping).	13	38	28	9	6	7	54	32	9		+
6. Use the results of standardized reading achievement tests as measured by the State Annual Testing Program.	50	41	6	3		50	32	11	7		−
7. Move students from one group to another (i.e., flexible grouping) as needs vary.	75	22				3	73	23	5		−
8. Group students on the basis of the group they were in last year.	6	22	66	3	3	14	36	36	11	2	+
9. Organize my reading program with the help of reading specialists.	34	31	22	9	3	10	40	25	15	10	−
10. Use parents, para-professionals and/or community resources.	44	41	13	6		25	57	9	9		0
11. Organize my reading program with the help of colleagues.	50	34	16			71	21	3	3	3	+

P = .5001 N.S.

Finally, an attempt was made to tie responses on the surveys directly to components of the READ curriculum. A sign test was applied to the items from the Principal and Teacher Surveys that related directly to components of READ. The overall hypothesis pertaining to the use of practices and procedures related to the READ curriculum was tested by using this procedure.

Hypothesis 2. Satisfaction of principals, teachers, parents, and students with the reading program. The impact of the program on the attitudes of principals, teachers, parents, and students was obtained from surveys designed by the Reading Curriculum Committee in 1979 and administered in 1979 and 1983. The same principals and teachers who responded to items related to the description of the reading program also responded to items 63 to 67 of the Principal and Teacher Surveys in Reading relating to attitudes about the reading program. A 5% random sample of students stratified by levels K to 3, 4 to 6, 7 to 9, and 10 to 12 was obtained using a computer generated list. These students received the Student Survey in Reading. The parents of these same students were sent the Parent Surveys in Reading. All surveys were distributed by school principals in 1979 and by members of the Reading Curriculum Committee in 1983. The returned surveys were processed by the statistical center at the University of North Carolina at Greensboro using SPSS (1979, 1982) programs to obtain the frequencies and percentages presented in this report.

The sign test, described earlier, was used to determine significant differences in clusters of items related to satisfaction with the reading program. These items are described as satisfaction because

they showed attitudes toward elements of the reading program. Table 7 contains a list of those items and indicates the survey on which the item was listed. Some of the items tapped satisfaction directly such as, "Are you satisfied with the reading program in your school?" Other items related to satisfaction only indirectly or through inferences. For example, more positive responses to item number eleven on the Parent Survey in Reading, "Do you think reading is an important skill?" implied more parental satisfaction through more positive responses to the importance of reading.

Table 7

Items from Principal, Teacher, Parent, Student K to 3,

and Student 4 to 12 Surveys that Relate to Satisfaction

with the Reading Program

Survey	Items
Principal and Teacher	64. How successful is your school's reading program on the whole?
	65. How satisfied are you with the reading skills and habits of the students in your classes?
	66. How effective do you find the present reporting system for record keeping of reading skills as students move from grade to grade?
	67. How effective do you find the present reporting system in reading for school to school transfer?
Parent	1. Do you feel that your child is making satisfactory progress in reading?

Table 7 (continued)

Survey	Items

2. Are you satisfied with the reading program in your child's school?

3. Would you feel comfortable discussing any reading problems your child might have with his or her teachers?

4. Does the school let you know about your child's progress (by report card or teacher conferences) in reading as often as you would like?

5. Has the school let you know your child's reading grade level?

6. Do you feel that your child enjoys learning to read   or improving his or her reading skills?

7. Does your child think reading is an important skill?

8. Do you know what type of reading activities your child does in school?

9. Do your child's reading teachers provide additional materials for him or her?

10. Does your child read at home?

11. Do you think reading is an important skill?

12. Does your child's reading teacher encourage your child to read at home?

Table 7 (continued)

Survey		Items
Student	K-3	1. Do you like what your teacher plans for you to read every day?
		2. Are you learning to read?
		3. Do you like to read?
		4. Do you read every day in school?
		5. Do you read at home?
Student	4-12	1. Do you know which activities you do at school that are supposed to help you read better?
		2. Do you read at home as part of homework?
		3. Do you read at home for fun?
		4. Does your school prove you with interesting materials to read?
		5. Does your school and/or teacher let you know about your progress in reading as often as you like?
		6. Do you ask your reading/language arts teachers for help if you have problems with reading?
		7. Does your teacher give you help with reading when you need it?
		8. Are you satisfied with the help you are receiving?
		9. Do you feel that your reading has improved this school year?
		10. Do you enjoy your reading activities at school?

Table 7 (continued)

Survey	Items
	11. Do you think reading is important?
	12. Are you happy with the way you read?

Although the overall hypothesis was related to satisfaction with the reading program, the sign test was applied separately for principals', teachers', parents', students' K to 3 and students' 4 to 12 satisfaction with the reading program. Responses made in 1979 were compared to those made in 1983. Responses that indicated yes or sometimes were combined and used as positive responses. Each item was assigned a plus if the percentage of positive responses was greater in 1983 and a minus if the percentage of positive responses was smaller in 1979. For example, items 1 to 5 on the Student K to 3 Survey in Reading were clustered in order to determine whether students' attitudes about the reading program were more positive (See Table 8). The probability of obtaining the two pluses with five items was .5000 which was not significant at the .05 level. Therefore, the directional hypothesis was not retained for K to 3 students' satisfaction with the reading program. The other clusters of items from the Principal, Teacher, Parent, and Student 4 to 12 Surveys were analyzed in a similar manner.

Table 8

Example of the Sign Test Applied to the

Student Survey in Reading K to 3

This survey is part of an attempt to describe the present reading program of the Guilford County Schools. It is to be completed by the student in the presence of an adult. The adult will read the questions to the student and record the student's answer.

DIRECTIONS: Choose the appropriate answer for the following statements and mark the correct box.

    A. Yes
    B. Sometimes
    C. No
    D. Don't Know
    E. Not applicable.

    1983   1979 = +
    1979   1983 = -
    No signs are attached to ties
    NS = Not Significant

	Percentage of Responses 1979 n=504					1983 n=417					Direction Of Difference
	A	B	C	D	E	A	B	C	D	E	
1. Do you like what your teacher plans for you to read every day?	85	11	2	1	0	84	12	3	1	0	0
2. Are you learning to read?	95	2	1	1	0	94	2	3	1	0	-
3. Do you like to read?	86	9	3	1	0	84	11	3	1	0	0
4. Do you read every day in school?	71	16	13	0	0	71	17	12	1	0	+
5. Do you read at home?	61	31	8	0	0	65	28	7	0	0	+

P = .5000 NS

Therefore, the sign test was applied five times for subhypotheses related to the overall hypothesis for satisfaction with the reading program. For these analyses, probabilities equal to or less than .05 for the obtained number of minuses were sufficient to retain the directional hypotheses. The results from these analyses are discussed in Chapter IV. The Principal, Teacher, Parent, Student K to 3 and Student 4 to 12 Surveys are found in Appendices A to E respectively.

The overall hypothesis related to satisfaction with the reading program was tested by applying the sign test to all items from all surveys that could be tied to the READ curriculum. Thus from the total analyses, overall satisfaction with the reading program as well as satisfaction about various components of READ were determined. As stated earlier, the surveys were planned before the READ curriculum was planned and implemented; therefore, there is not a one-to-one relationship between items on the surveys and components of the READ curriculum.

Hypothesis 3. Reading Achievement. Data related to the impact of the program on the reading achievement of the students were collected from archival data on file in the Office of Research, Planning, and Evaluation. These data consisted of standardized test results for 1979 and 1982 from the North Carolina Annual Testing Program. The 1983 reading scores had not been processed at the time of this study.

An analysis of covariance, controlling for variables related to school effects, was utilized to compare mean school system reading achievement scores of third and sixth graders in 1979 to mean school

system reading achievement scores of third and sixth graders,
respectively, in 1982. This was done to determine whether the overall
mean school system reading achievement was significantly higher in
1982. The mean school system reading achievement scores and data
on the control variables were obtained from mean school data from
21 schools.   The school was used as the unit of analysis
not only for practical reasons such as time and cost of data analysis,
but also because school data for this school system is probably more
consistent from year to year than student data. The control variables
used in the analysis of covariance were (a) mean school IQ as measured
by the Short Form Test of Academic Aptitude (SFTAA) which is routinely
administered as a part of the North Carolina Annual Testing Program;
(b) percentage of students in the federally funded Chapter I remedial
reading program, which is a rough estimate of the socioeconomic status
of the school since, by mandate, this program only exists in schools
with a high percentage of students receiving free lunches; (c) percentage of
parents who have not completed high school, which is an indication of
the educational level of the parents; and (d) percentage of minority
students, which may affect school test scores since standardized tests
historically have not always measured accurately the achievement of
this body of students.

Data related to these variables were initially obtained from
classroom teachers who must submit these identifying data as a part
of the North Carolina annual testing program. The classroom teachers'
data were summarized and an average was obtained for each school on
these variables.

Mean school system vocabulary, comprehension, and total reading scores were compared for both grades 3 and 6. The F ratios from these analyses were used to determine differences between means that were significant at the .05 level. Therefore, there were six subhypotheses related to the overall hypothesis for student achievement in reading. The directional hypotheses were retained for F ratios that were significant and the corresponding unstated null hypotheses were rejected.

Hypothesis 4. Validating the causal inference that program effects were due to READ. A metaevaluation technique was employed in order to determine whether the changes in the program were due to the READ curriculum and not to extraneous factors. Cook and Gruder (1978) used the term metaevaluation to refer to the evaluation of summative evaluation studies. The purpose of metaevaluation was described as improving the technical quality of an evaluation. While Cook and Gruder focused on metaevaluation for improving technical aspects of evaluation, it can serve another function. In this study, metaevaluation was used to validate or support the inference that the READ curriculum was the cause of any observed program effects. As noted in the earlier discussion of the READ impact evaluation, the design is quasi-experimental. While students were retested after program implementation, the lack of an experimental control group resulted in weak support for a causal inference. The procedures used in this phase of the study were designed to meet this need.

Cook and Gruder (1978) described several metaevaluation models which they generated from an analogy with a three-factor analysis of

variance. These models were presented in Chapter II. Of the seven
models presented by Cook and Gruder, model number seven, multiple
independent replications, was deemed more appropriate for this
study.

In a large school system, it is impossible to present detailed
curriculum and evaluation plans to fit each school. The needs of the
students are diverse and the amount of parental support varies. These
and other factors make it difficult for a school district to assess
program effect and determine that a program innovation produced the
results. In particular, asking only one school in a district to assess
the effects of curriculum intervention would in all likelihood produce
biased estimates for the district. The use of multiple replications
for the metaevaluation was considered valuable not only for reducing
biased results but also for strengthening the link between program
effects and the causal inference that outcomes were due to implementing
the READ curriculum. Further, the metaevaluation was to be conducted
at two levels, at the system level using the system-wide Reading
Curriculum Committee already established, and at the school level
using a random sample of schools.

It was decided that a random sample of at least six schools (14%)
would produce results from which inferences could be made. Since the
reading program affects all schools, the names of all 44 schools were
placed in a box. Six were drawn to be used in the study. The
representatives of those schools serving on the Reading Curriculum
Committee were asked to carry out the set of procedures. The general
procedures were to use two rounds of surveys conducted in parallel

for the system and school levels. The first round of surveys was used
to identify practices and procedures that were practiced more
frequently since implementing READ. The first round survey was called
Impact Questionnaire #1. The same six teachers in each school who
had responded to the Teacher   Survey in Reading were asked to rate
each item on Impact Questionnaire #1 according to how much the
practice had changed since implementing the READ curriculum. The
Reading Curriculum Committee completed the same survey. The original
Teacher   Survey in Reading was revised to accomodate the ratings.
The range of possible ratings were:

+2 This is practiced much more often since implementing
     the READ curriculum,

+1 This is practiced a little more often,

 0 This practice has not changed,

-1 This is practiced a little less often since implementing
     the READ curriculum, and

- 2 This is practiced much less often since implementing the
      READ curriculum.

Appendix F contains a copy of Impact Questionnaire #1. The
metaevaluation is summarized in Figure 1.

```
┌─────────────────────┐              ┌──────────────────────────┐
│ SYSTEMWIDE READING  │              │ 6 RANDOMLY SELECTED SCHOOLS│
│ Curriculum Committee│              │  n = 6 Teachers per school │
│      N = 44         │              │          N = 36            │
│ one representative  │              └──────────────────────────┘
│  from each school   │
└─────────────────────┘
```

```
┌────────────────────────────────────────────────┐
│       Round 1 - Impact Questionnaire #1 - Rating│
│       69 items using ratings from +2 to -2      │
└────────────────────────────────────────────────┘
```

```
┌───────────────────┐              ┌──────────────────────────┐
│ Items Receiving   │              │  Items Receiving a Rating │
│ A Rating of at    │              │      of at Least 6        │
│   Least 44        │              └──────────────────────────┘
└───────────────────┘
```

```
┌──────────────────────────────────────────────────┐
│        Round 2 - Impact Questionnaire #2           │
│  Select 10 of these items that relate most closely to│
│   the READ Curriculum.  Rank these itmes from 1-10 │
│          (1 = most closely related)                │
└──────────────────────────────────────────────────┘
```

```
┌──────────────────────────────────────────┐
│          Comparing the lists using        │
│    Kendall's Coefficient of Concordance   │
└──────────────────────────────────────────┘
```

```
┌──────────────────────┐
│    Items appearing    │
│      on each list     │
└──────────────────────┘
```

```
┌──────────────────────────────────────────┐
│    Effects of the Reading Program         │
│    Attributable to the Curriculum         │
└──────────────────────────────────────────┘
```

FIGURE 1 Metaevaluation Model
Independent Multiple Replications

The ratings were tallied for each item. Those items that received a total rating of at least 6 (or n x 1) in each school were listed on Impact Questionnaire #2 for the school. For the Reading Curriculum Committee, those items that received a rating of at least 44 (n x 1) were placed on Impact Questionnaire #2. Each study was conducted independently; therefore, there were seven questionnaires for round two, one for the Reading Curriculum Committee and six for the randomly selected schools.

The same respondents were then asked to select the top ten items on Impact Questionnaire #2 and rank them on a scale from 1 to 10 on the basis of how closely teachers felt that they related to the READ curriculum, where 1 indicated most closely related. The rankings were then assigned weights 1 = 10, 2 = 9, 3 = 8, and so on, which were summed for each item in order to determine the 10 items receiving the highest points. The items within a replication that received the highest points were given a new rank of 1, 2, 3, and so on. Each replication had a final list of ten items ranked from 1 to 10 on the basis of a closeness to the READ curriculum.

The following example illustrates how the two rounds of surveys operated. Item 12 on the Teacher Survey, "Use the school system's curriculum guide in reading as the source of reading program goals," was indicated as being practiced more frequently since implementing READ. Item 12 was placed on the round two survey along with other items identified in the same way. On the second survey, the teachers were asked to rank the items that were practiced more frequently on the basis of how closely the related to READ. The purpose of the second

round of surveys was to eliminate items that were practiced more
frequently since implementing READ but which may not have been tied
to the READ curriculum. Item 12, use of the school system's curricu-
lum guide in reading as a source of reading program goals, received
ranks ranging from 2 to 7 on the second round of surveys. Since this
item was ranked among the top 10 items on all replications, it was
considered an impact statement related to the READ curriculum. This
procedure added validity and credibility to the impact statement that
since the implementation of READ, more teachers use the system's
curriculum guide in reading as a source of reading program goals.

Hypothesis 4 was tested by using Kendall's coefficient of
concordance (Daniel, 1978). Kendall's coefficient was computed in
order to determine if there was significant agreement among the seven
lists. Significant agreement among the lists was not considered
undisputed evidence of the effects of the READ curriculum; therefore,
the additional criterion was added that an item must appear on the
final list of each of the replications before it would be considered
an undisputed impact of the READ curriculum.

The use of multiple replications in the adaptation of a
metaevaluation was based on the Reading Curriculum Committee and
randomly selected schools independently giving perceptual data about
factors attributable to the READ curriculum. These procedures were
viewed as a way to add validity and reliability to the study. Validity
in this context was defined as the degree to which the stated program
effects which were practiced more frequently since implementing READ
were judged as being impacts of READ. Reliability was defined as the

consistency of reported effects among the seven generated lists of impact statements.

CHAPTER IV

FINDINGS AND DISCUSSION

The major findings from the study will be presented as they relate
to the four hypotheses. The responses to the pretest and posttest
surveys were analyzed to test hypothesis 1 related to describing the
program and hypothesis 2 related to satisfaction with the reading
program. Hypothesis 3 was tested by analyzing third and sixth grade
reading achievement scores from the North Carolina Annual Testing
Program secured from the Office of Research, Planning, and Evaluation
of the Guilford County School System. Hypothesis 4 was tested by having
multiple, independent replications of two rounds of impact question-
naires. Since these were directional hypotheses, each was subjected to
a one-tailed test. For significant differences at the .05 level, the
directional hypotheses were accepted and the corresponding null hypo-
theses were rejected. For nonsignificant differences, the directional
hypotheses were rejected; the corresponding unstated null hypotheses
were not rejected.

Hypothesis 1

Hypothesis 1 stated, "Principals and teachers will report a
greater use of practices and procedures related to the READ curriculum
in 1983 than they did in 1979." This hypothesis was tested at the .05
level.

The sign test was applied to categories of items from the Principal and Teacher surveys in order to test this hypothesis. The items in each descriptive category upon which the sign test was applied were items 1 to 11 for Organizing for Reading Instruction, 12 to 15 for Concerning Goals and Objectives, 16 to 20 for Planning for Skills Instruction, 21 to 29 for Instructing Students in Reading, 30 to 39 for Methods of Teaching, 40 to 46 for Teacher Evaluation Practices, and 47 to 63 for Teacher's and Principal's Knowledge and Practices. The data for the sign tests are summarized in Table 9. The Principal and Teacher Surveys in Reading along with the 1979 and 1983 data which were used to test hypothesis 1 may be found in Appendixes A and B, respectively.

Table 9

Summary of Probability Values for Sign Tests for Teachers' and Principals' Description of the Reading Program in 1979 and 1983

	Teachers		Principals	
	Differences in Responses from 1979 to 1983			
	1979 n=641	1983 n=270	1979 n=32	1983 n=44
	P =		P =	
Organizing Students for Reading Instruction	*.0059		.5001	
Concerning Goals and Objectives	.0625		.0625	
Planning for Skills Instruction	.1274		.4999	
Instructing Students in Reading	*.0076		.2539	
Methods of Teaching	*.0351		.2539	
Teacher Evaluation Practices	*.0156		.1094	
Teacher's and Principal's Knowledge and Practices	*.0002		*.0245	

*Significant at p $\leq$ .05

On the Teacher Survey, changes in the description of the reading program were found in the following areas: Organizing Students for Reading Instruction, Instructing Students in Reading, Methods of Teaching, Teacher Evaluation Practices and Teacher's Knowledge and Practices. The only area found to be significantly different on the Principal Survey was in Principal's Knowledge and Practices.

Further analysis of the various sections on the Principal and eacher Surveys indicated that the changes most related to the READ curriculum were in the section under Teacher's and Principal's Knowledge and Practices.

The largest shift in responses favoring 1983 was in knowledge and use of the school system's curricular guide, a component of READ. In 1979, 82% of the teachers indicated knowledge of the guide. In 1983, 90% of the teachers indicated such knowledge. In 1979, 70% of the teachers and 85% of the principals indicated that the guide was used compared to 87% of the teachers and 90% of the principals in 1983.

In 1982, 86% of the teachers indicated that the basal reader was used for instruction as opposed to 78% in 1979. The systemwide adoption of a basal reader and the use of accompanying periodic and cumulative tests (components of READ) may have contributed to this shift in responses.

It should be noted here that all of the surveys were planned in 1979 before the READ curriculum was planned and implemented; therefore,

it was difficult to tie responses on the surveys directly to READ.

However, Table 10 illustrates that when the sign test was applied,

those items on the surveys that could be linked to the READ curriculum

did show more positive responses for 1983. Therefore, with $\underline{p} < .0095$

for principals and $p < .0002$ for teachers, the overall hypothesis related

to the description of the reading program was thus retained.

Table 10

READ Program Components and Matching Descriptive Item on the

Principal and Teacher Surveys with Appropriate

Sign After Applying the Sign Test

READ Components	Survey Item	Principals' and Teachers' Sign	
Curricular Guide with K to 12 reading objectives	12. Use of the school system's curricular guide as a source of reading program goals.	+	+
	52. I have knowledge of the curricular guide of our school system.	+	+
	53. I use this guide.	+	+
Lead Basal Reader and Co-basal Reader Series	1. Divide students into more than one group on the basis of reading ability.	+	+
	15. Use the basal reader as the source of program goals.	+	+
	18. Use the basal reading series to determine the sequence in which reading skills are taught.	+	+
	30. Use the basal approach.	-	+

Table 10 (continued)

READ Components		Survey Item	Principals' and Teachers' Sign	
	45.	Consider to what degree a student's textbooks or required reading are matched to his/her reading level.	+	+
Checklist of Reading Skills as reading record to be passed from grade to grade and from school to school.	14.	Maintain a record keeping system to keep track of individual progress toward specific objectives.	+	+
Reporting Form for Parents				
Periodic and Cumulative Tests	43.	Compare achievement to behavioral objectives by means of criterion-referenced tests.	+	+
Procedure for Resource File	19.	Select practice activities that match their instructional objectives.	+	+
	21.	Teach reading through nonbasal materials.	−	+
	23.	Use a file of workbook pages and exercises classified by skill and level.	+	+
	26.	Prepare a directory of commercial reading materials available within my school.	+	+
	28.	Code reading materials to reading objectives.	+	0
Correction of READ objectives with CAT and PRI	6.	Use the results of standardized reading achievement tests as measured by the State Annual Testing Program.	−	+
		n = 16	p $\leq$ .009	p $\leq$ .002

Hypothesis 2

Hypothesis 2 stated, "Principals, teachers, parents, and students will show greater satisfaction with the reading program in 1983 than they did in 1979." This hypothesis was tested at the .05 level.

The sign test was applied to clusters of items on the Principal, Teacher, Parent, Student K to 3, and Student 4 to 12 Surveys in Reading to determine the attitudes of these various groups toward the reading program. The sign test was applied to items 64 to 67 on the Principal and Teacher Survey, to items 1 to 13 on the Parent Survey, to items 1 to 5 on the K to 3 Student Survey, and items 1 to 13 on the 4 to 12 Student Survey. The results from the application of the sign tests are shown in Table 11.

Table 11

Summary of Probability Values of Sign Test for Satisfaction

with the Reading Program

Survey	n=		Differences in Responses from 1979 to 1983 P=
	1979	1983	
K-3 Student Survey in Reading	504	417	.5000
4-12 Student Survey in Reading	645	714	.1124
Parent Survey in Reading	754	918	*.0059 < .05
Teacher Survey in Reading	641	270	.99
Principal Survey in Reading	32	44	.7723

*Significant at $p \leqslant .05$

A significant difference was found in the satisfaction level of parents with respect to the reading program. The hypothesis was retained for parent satisfaction with the reading program; the corresponding unstated null hypothesis was rejected. The directional hypothesis was rejected for principals, teachers, and students because the cluster of responses for 1983 did not differ significantly from the cluster of responses given in 1979 with respect to satisfaction with the reading program. The corresponding null hypotheses were not rejected.

Since there was not a perfect match between READ and items on the surveys, an attempt was made to assess the satisfaction of principals, teachers, parents, and students with the READ curriculum by looking at various survey items that related to components of READ. Regarding the Checklist of Reading Skills, principals and teachers made positive responses about the effectiveness of the reporting system from school to school. More teachers responded positively to the reporting system from grade to grade, but fewer principals showed satisfaction from grade to grade. With respect to the Reporting Form for Parents, more parents and 4 to 12 students indicated that they received information about progress in reading. When the sign test was applied to these items, the probability of .0625 was not significant. Table 12 lists these items and the results of the sign test.

Table 12

READ Components and Matching Satisfaction Items on the Principal,

Teacher, Parent, Student K to 3, and Student

4 to 12 Surveys with Appropriate Sign After

Applying the Sign Test

READ Component	Survey Item	Sign
Checklist of Reading Skills as a reading record to be passed from grade to grade and from school to school.	Principal Survey	
	66. How effective do you find the reporting system from grade to grade	-
	67. How effective do you find the reporting system from school to school.	+
	Teacher Survey	
	66. How effective do you find the reporting system from grade to grade.	+
	67. How effective do you find the reporting system from school to school.	+
Reporting Form for Parents	Parent Survey	
	4. Does your school let you know your child's progress in reading as often as you would like?	+
	5. Has the school let you know your child's reading level?	+

Table 12 (continued)

READ Component	Survey Item	Sign
	Student 4-12	
	5. Does your school or teacher let you know your progress in reading as often as you like?	+

$p \leq .0625$

In order to test the overall hypothesis related to satisfaction of principals, teachers, parents, and students with the reading program, all items from the surveys that tapped satisfaction directly were clustered. More positive responses were found in 1983 for eight of the eleven items. The probability of this occurrence was .0461 which was significant at the .05 level. Therefore, hypothesis 2 was retained. Table 13 summarizes these results.

Table 13

Results of Sign Test on Items that Tap Satisfaction with the

Reading Program in a Direct Way

Survey	Item	Sign
Principals	64. How successful is the reading program in your school?	-

Table 13 (continued)

Survey	Item	Sign
	65. How satisfied are you with the reading skills and habits of the students in your classes?	+
Teachers	64. How successful is the reading program in your school?	+
	65. How satisfied are you with the reading skills and habits of the students in your classes?	+
Parents	1. Do you feel that your child is making satisfactory progress in reading?	+
	2. Are you satisfied with the reading program in your child's school?	+
	6. Do you feel that your child enjoys learning to read and/or improving his/her reading skills.	+
Students K to 3	1. Do you like what your teacher plans for you to read everyday?	0
	2. Are you learning to read?	-
	3. Do you like to read?	0
Students 4 to 12	8. Are you satisfied with the help you are receiving in reading?	+
	9. Do you feel that your reading has improved this year?	+

Table 13 (continued)

Survey	Item	Sign
	10. Do you enjoy your reading activities at school?	-

$p \leq .0461$

Hypothesis 3

Hypothesis 3 stated, "Students in grades three and six will, on the average, attain higher reading achievement scores on the CAT in 1982 than students in grades three and six attained in 1979." This hypothesis was tested at the .05 level.

Hypothesis 3 was tested using an analysis of covariance (ANCOVA) to determine if there were significant differences in the mean school system CAT reading scores of third graders in 1979 and 1982 and sixth graders for the same years. Mean school system vocabulary, comprehension, and total reading scores were compared within the third and sixth grades. ANCOVA enables the researcher to identify and take into account sources or variance due to concomitant variables thereby providing greater control. According to Pedhazur (1982), if the variability of a given concomitant variable is relatively large, and this variable is correlated with the dependent measure, it is possible to use the subjects' scores on the concomitant variable as a covariate. An adjustment for the covariate will lead to a reduction in the error term, and consequently to a more sensitive analysis. The covariates for these analyses were averaged to obtain crude estimates of mean school IQ as measured by the SFTAA, percentage of students in Chapter I remedial reading programs, percentage of parents who had not completed high school, and percentage of minority

students. These estimates were treated as interval scores for those
analyses. Additional information about the covariates was presented
in Chapter 2.

Unadjusted means and standard deviations of the reading scores and
the covariates which were used in these analyses are shown in Table 14
for the third grade and in Table 15 for the sixth grade. There were
increases in mean scale scores for both grades in vocabulary,
comprehension, and total reading from 1979 to 1982. The mean vocabulary
score of 405.86 for the third grade was below the national norm of 408
in 1979. In 1982, the mean vocabulary score was 419.33, well above the
national norm of 411. The mean scale score for comprehension of 422.62
in 1979 was three scale score points above the national norm of 419 as
compared to 427 in 1982 which was well above the national norm of 422.
This pattern was consistent also for total third grade reading, sixth
grade vocabulary, sixth grade comprehension, and sixth grade total
reading.

The standard deviations of the third grade scores indicated a
decrease in variability from 1979 to 1982. For vocabulary, the
standard deviation decreased from 13.69 to 9.37, for comprehension the
change was from 12.59 in 1979 to 437.48 in 1982, and for total reading,
the standard deviation decreased from 15.78 in 1972 to 12.26 in 1982.
Conversely, the standard deviations of the sixth grade indicated a greater
spread from 1979 to 1982. From 1979 to 1982 the changes for the sixth
grade were 14.89 to 16.92 for vocabulary, 13.61 to 15.34 for comprehension,
and 15.02 to 16.92 for total reading. This phenomenon did not appear to
be a measurement error. Rather, it appeared to be a function of the
reading program. The reading program may in fact limit diversity in the

Table 14

School Means and Standard Deviations on Reading Test

Scores and Covariate Measures for Grade Three

in 1979 and 1982

	X	SD	National Student Norms Spring Testing	
			X	SD
1979 Vocabulary	*405.86	13.69	408	52.3
1982 Vocabulary	419.33	9.37	411	
1979 Comprehension	422.62	12.59	419	57.2
1982 Comprehension	437.48	10.42	422	
1979 Total reading	403.05	15.78	401	59.0
1982 Total reading	420.43	12.26		
1979 % Minority	19.86	9.00		
1982 % Minority	17.00	14.09		
1979 Mean School IQ	102.57	4.00		
1982 Mean School IQ	101.48	4.25		
1979 % Chapter I	15.66	16.00		
1982 % Chapter I	16.62	13.07		
1979 % Parents Not Finished High School	13.52	11.25		
1982 % Parents Not Finished High School	20.04	9.75		

*CAT Scale Scores

N = 21 Schools

Table 15

School Means and Standard Deviations on Reading Test

Scores and Covariate Measures for Grade Six

in 1979 and 1982

	X	SD	National Student Norms Spring Testing	
			X	SD
1979 Vocabulary	503.52	14.89	498	67.7
1982 Vocabulary	511.48	16.92	499	
1979 Comprehension	522.52	13.61	510	71.2
1982 Comprehension	536.91	15.34	512	
1979 Total Reading	509.00	15.02	500	70.0
1982 Total Reading	520.48	16.92	500	
1979 % Minority	17.43	10.25		
1982 % Minority	16.95	9.76		
1979 Mean School IQ	101.19	3.00		
1982 Mean School IQ	100.57	3.94		
1979 % Chapter I	11.76	8.00		
1982 % Chapter I	10.00	10.05		
1979 % Parents Not Finished High School	20.62	14.50		
1982 % Parents Not Finished High School	22.90	13.43		

*CAT Scale Scores

N = 21 Schools

early grades by pacing the students more slowly. On the other hand,
diversity is increased at the upper levels where interests and aptitude
have more bearing. That there was less variability within the local
scores than were illustrated by the national norms is understandable
when one considers that the local mean scores were based on mean school
data while the national norms were based on individual data.

The 1979 mean system reading scores listed in Table 14 for the
third grade and in Table 15 for the sixth grade were obtained by
averaging mean school data from 21 schools. This procedure violated
one of the assumptions of ANCOVA, random assignment of individuals to
treatments (Elashoff, 1969). This was not considered a serious violation
in view of the similarity of school means on the control variables for
both years.

Correlation coefficients showing the strength and direction of the
relationship between reading scores and the covariates indicated strong
positive correlations between IQ and reading achievement and negative
correlations between percentage of students in Chapter I, percentage of parents
who had not completed high school, and percentage of minority enrollment.
These correlation coefficients supported the inference that each of the
covariates should be included in the ANCOVA. The correlation coefficient
between third grade total reading and the covariates ranged from -.476
for percentage of minority to .902 for mean school IQ. These correlation
coefficients are presented in Table 16 for the third grade and in
Table 17 for the sixth grade. Data for 1979 were used to obtain these
correlations.

Table 16

Correlation Coefficients for 1979 Third Grade Reading

Scores and Covariates

	Vocabulary	Comprehension	Total Reading	% Minority	Mean School IQ	% Chapter I	% Parents Not Completed High School
Vocabulary	1.000	.995	.997	-.479	.936	-.760	-.448
Comprehension		1.000	.997	-.426	.887	-.762	-.503
Total Reading			1.000	-.476	.902	-.788	-.497
% Minority				1.000	.472	.559	.257
Mean School IQ					1.000	-.773	-.593
% Chapter I						1.000	.573
% Parents Not Completed High School							1.000

Table 17

Correlation Coefficients for 1979 Sixth Grade Reading

Scores and Covariates

	Vocabulary	Comprehension	Total Reading	% Minority	Mean School IQ	% Chapter I	% Parents Not Completed High School
Vocabulary	1.000	.995	.997	-.420	.947	-.752	-.668
Comprehension		1.000	.997	-.472	.918	-.664	-.530
Total Reading			1.000	-.462	.951	-.718	-.602
% Minority				1.000	.445	.521	.028
Mean School IQ					1.000	-.793	-.678
% Chapter I						1.000	.600
% Parents Not Completed High School							1.000

The model for the ANCOVA took the basic form:

$$\hat{Y}_{ij} = \overline{Y} + \alpha_j + B_1(X_{1ij} - \overline{X}_1) + B_2(X_{2ij} - \overline{X}_2) + B_3(X_{3ij} - \overline{X}_3) + B_4(X_{4ij} - \overline{X}_4) + E$$

Where $\hat{Y}_{ij}$ = the adjusted means for any school i.

$\overline{Y}$ = grand mean on the independent variables

α_j = treatment (year) effect

B_1 = the regression coefficient due to % minority (Y on x_1)

B_2 = the regression coefficient due to mean school IQ (Y on x_2)

B_3 = the regression coefficient due to Chapter I (Y on x_3)

B_4 = the regression coefficient due to parents not finished high

    school (Y on x_4)

X_{1ij} = score on the covariate for % minority for school ij

X_{2ij} = mean school IQ of school ij

X_{3ij} = score of the covariate for % Chapter I for school ij

X_{4ij} = % parents not finished high school for school ij

$\overline{X}_{1ij}$ = overall mean for % minority

$\overline{X}_{2ij}$ = overall mean for IQ

$\overline{X}_{3ij}$ = overall mean for % Chapter I

$\overline{X}_{4ij}$ = overall mean for % parents not finished high school

E = random error

Pedhazur (1982) and Cook and Campbell (1979) discussed the ANCOVA model in more detail. The following equations illustrates the ANCOVA model to the third grade total reading score to estimate the system mean adjusted for the covariates:

$$\hat{Y}_{ij} = \overline{Y} + \alpha + B_1 (X_1 - \overline{X}_1) + B_2(X_2 - \overline{X}_2) + B_3(X_3 - \overline{X}_3) + B_4(X_4 - \overline{X}_4) + E$$

$$= \overline{Y} + \alpha + .109(-1.43) + 2.53(-.545) + -.344(1.775) + .319(3.26)$$

1979 $\quad \hat{Y} = 403.05 + \quad + (.15655) + (1.3799) + (.6112) + (-1.0427) + E$

$\qquad \hat{Y} = 404.15$

1982 $\quad \hat{Y} = 420.43 + \alpha + (-.15655) + (-1.37992) + (-.6112) + (1.0427) + E$

$\qquad \hat{Y} = 419.32$

Covariates		1979	1982	$\overline{X}i79+82$	1979 X-$\overline{X}$	1982 X-$\overline{X}$	$B_i(X_i-\overline{X}_i)$
% Minority	X_1	19.86	17.00	$\overline{X}_1$= 18.43	1.43	-1.43	$\pm$ .15
Mean School							
IQ	X_2	102.57	101.48	$\overline{X}_2$=102.03	.54	- .545	$\pm$ 1.37
% Chapter I	X_3	13.07	16.62	$\overline{X}_3$= 14.84	-1.77	1.775	$\pm$ .61
% Parents Not							
Finished							
High School	X_4	13.52	20.00	$\overline{X}_4$= 16.78	-3.26	3.26	$\pm$ 1.04

These procedures produced data that approximated that data from a
computer analysis using the SPSS package. These data are illustrative.
Data from the computer analyses were used to formally test
hypothesis three.

When 1982 reading scores were compared to 1979 scores, significant
differences were found for both third and sixth grade vocabulary,
comprehension, and total reading. The $\underline{F}$ ratios for the third grade were
vocabulary 47.74, comprehension 44.45, and total reading 39.87, all
significant at the .05 level. For the sixth grade the $\underline{F}$ ratios were 7.19
for vocabulary, 26.78 for comprehension, and 16.53 for total reading
which were also significant. In addition, the sum of squares due to
regression were significant for all comparisons. This was an indication

that the covariates did contribute to increasing the precision of
the analyses. Precision is increased when the initial difference on a
covariate reduces the error term. This increase in precision occurs
only when a covariate is significantly related to the dependent variable.
For example, if it is known that individuals or schools differ in mental
ability, and it is known that this variable is related to performance on
the dependent variable, this source of variability (due to IQ) may be
indirectly controlled by the use of ANCOVA (Pedhazur, 1982). However,
when several covariates are used, each is tested for the significance
of its contribution to the dependent variable. In addition, when the
correlation coefficients showing the direction and strength of the
relationship between each covariate and the dependent variables are
similar over all treatments, it can be intuitively concluded that the
regression coefficients used in the analyses are homogeneous. The
results of the ANCOVA are presented in Tables 18 to 23.

Tables 18 to 23 also present the t tests to determine whether
the use of each covariate added significantly to the proportion of
variance accounted for in the dependent variable. Mean school IQ was
the only covariate that met the test of significance consistently over
all comparisons. Percentage of minority was not significant for any
comparison and therefore is apparently not a useful covariate for similar
studies within this school system. The t tests for percentage of students
in Chapter I programs and percentage of parents who had not finished high
school were significant for third grade total reading and third grade
vocabulary. They were not significant for the other comparisons. For
the sixth grade comparisons, the t tests indicated that mean school IQ
was the only significant covariate.

Table 18

Analysis of Covariance Between Third Grade Total Reading

Scores in 1979 and 1982

Source of Variation	Sum of Squares	df	Mean Squares	F	Significance of F
Within Cells	2108.58	36	58.88		
Regression	5883.52	4	1470.88	25.11	.000*
Constant	723.37	1	723.37	12.35	.001*
Year	2334.99	1	2334.99	39.87	.000*

Covariates	B	Beta	Standard Error	t	Significance of t
% Minority	.109	.09	.128	.850	.401
Mean School IQ	2.532	.796	.399	6.340	.000*
% Chapter I	-.344	-.355	.124	-2.757	.009*
% Parents Not Finished High School	.319	.242	.153	2.084	.044*

*Significant at p ∠ .05

Table 19

Analysis of Covariance Between Third Grade Vocabulary

Scores in 1979 and 1982

Source of Variation	Sum of Squares	df	Mean Squares	F	Significance of F
Within Cells	1144.81	36	31.80		
Regression	4362.42	4	1090.61	34.29	.000*
Constant	1089.12	1	1089.12	34.24	.000*
Year	1518.05	1	1518.06	47.74	.000*

Covariates	B	Beta	Standard Error	t	Significance of t
% Minority	.008	.008	.094	.085	.932
Mean School IQ	2.204	.835	.294	7.491	.000*
% Chapter I	-.248	-.308	.092	-2.697	.011*
% Parents Not Completed High School	.329	.300	.163	2.912	.006*

*Significant at p $<$ .05

Table 20

Analysis of Covariance Between Third Grade Comprehension

Scores in 1979 and 1982

Source of Variation	Sum of Squares	df	Mean Squares	F	Significance of F
Within Cells	1259.53	36	34.98		
Regression	4080.65	4	1020.16	29.15	.000*
Constant	1285.46	1	1285.46	36.74	.000*
Year	1555.01	1	1555.01	44.44	.000*

Covariates	B	Beta	Standard Error	t	Significance of t
% Minority	.040	.044	.099	.416	.684
Mean School IQ	2.226	.856	.308	7.212	.000*
% Chapter I	-.139	-.176	.096	-1.449	.156
% Parents Not Completed High School	.141	.130	.118	1.189	.242

*Significant at p $<$ .05

Table 21

Analysis of Covariance Between Sixth Grade Total Reading

Scores in 1979 and 1982

Source of Variation	Sum of Squares	df	Mean Squares	F	Significance of F
Within Cells	1553.49	36	43.15		
Regression	8683.75	4	2170.94	50.31	.000*
Constant	1354.59	1	1354.59	31.39	.000*
Year	713.35	1	713.35	16.53	.000*

Covariates	B	Beta	Standard Error	t	Significance of t
% Minority	-.109	-.075	.1198	-.914	.367
Mean School IQ	2.536	.677	.4433	5.721	.000*
% Chapter I	-.211	-.130	.1805	-1.173	.248
% Parents Not Finished High School	-.159	-.141	.1060	-1.502	.142

*Significant at p $<$ .05

Table 22

Analysis of Covariance Between Sixth Grade Vocabulary

Scores in 1979 and 1982

Source of Variation	Sum of Squares	df	Mean Squares	F	Significance of F
Within Cells	1056.465	36	29.35		
Regression	9102.02	4	2275.51	77.54	.000*
Constant	971.18	1	971.17	33.09	.000*
Year	210.95	1	210.94	7.18	.011*

Covariates	B	Beta	Standard Error	t	Significance of t
% Minority	-.028	-.071	.098	-1.040	.305
Mean School IQ	2.852	.7645	.365	7.802	.000*
% Chapter I	-.125	-.0773	.148	- .842	.405
% Parents Not Completed High School	-.136	-.1211	.087	-1.553	.128

*Significant at p $\leq$ .05

Table 23

Analysis of Covariance Between Sixth Grade Comprehension

Scores in 1979 and 1982

Source of Variation	Sum of Squares	df	Mean Squares	F	Significance of F
Within Cells	1831.82	36	50.88		
Regression	6777.21	4	1644.30	32.31	.000*
Constant	1819.49	1	1891.49	37.17	.000*
Year	1362.75	1	1362.75	26.78	.000*

Covariates	B	Beta	Standard Error	t	Significance of t
% Minority	-.073	-.056	.130	- .567	.574
Mean School IQ	2.195	.646	.481	4.560	.000
% Chapter I	-.208	-.141	.196	-1.062	-1.062
% Parents Not Finished High School	-.136	-.133	.115	-1.183	-1.183

*Significant at p ∠ .05

The ANCOVA made slight adjustments on the data for both years. Mean scores for 1979 were all adjusted upward. The 1979 third grade vocabulary mean was adjusted upward from 405.86 to 406.14. Similarly, 1979 sixth grade vocabulary was adjusted upward from 503.52 to 504.14. The 1982 mean scores were all adjusted downward. Third grade mean comprehension was adjusted from 437.48 down to 436.59. Likewise, sixth grade mean comprehension was adjusted from 536.91 down to 535.72. These data are presented in Table 24.

Table 24

Adjusted and Unadjusted System Level Means for

Grades Three and Six in

1979 and 1982

	Grade			
	3		6	
	Adjusted Mean	Unadjusted Mean	Adjusted Mean	Unadjusted Mean
1979 Vocabulary	406.14	405.86	505.14	503.52
1982 Vocabulary	419.05	419.33	509.86	511.48
1979 Comprehension	423.50	422.62	523.71	522.52
1982 Comprehension	436.59	437.48	535.72	536.91
1979 Total reading	403.73	403.05	510.39	509.00
1982 Total reading	419.75	420.43	519.08	520.00

Hypothesis 4

Hypothesis 4 stated, "Independent groups of school staff will agree on changes in the reading program that can be attributed to the READ curriculum." This hypothesis was tested at the .05 level.

Hypothesis 4 was tested by conducting two rounds of surveys, Impact Questionnaire #1 and Impact Questionnaire #2. On Impact Questionnaire #1, teachers rated the 69 items on the original Teacher Survey in Reading on the basis of having been practiced more frequently or less frequently since the implementation of the READ curriculum. Those items that received a total rating indicating that they were practiced more frequently since READ were placed on Impact Questionnaire #2. Teachers were asked to select the 10 items on Impact Questionnaire #2 that most closely related to READ and rank them from 1 to 10 on the basis of how closely they related to READ. Weights (points) were assigned to the ranks in order to compile the top ten items that were most closely related to READ.

Seven lists of impact statements were identified independently, one by the Reading Curriculum Committee and one for each of the randomly selected schools (N = 6). The lists of items which were derived from each replication are listed in Appendix H. Table 25 lists the ten impact statements identified by the Reading Curriculum Committee and the rankings given to those statements by the committee and the randomly selected schools. Kendall's Coefficient of Concordance was applied to the rankings on the seven lists to determine if there was any overlap in the lists and to test hypothesis 4. For the rankings on the seven lists, Kendall's Coefficient of Concordance was $W = 3.53$ which was

TABLE 25

Impact Statement and Rankings Given by the
Seven Replications on Impact
Questionnaire #2

		RANK					
IMPACT STATEMENTS	READING COMMITTEE	SCHOOL NUMBER					
		320	334	336	360	368	411
1 Teacher's knowledge of the school system's curricular guide has increased.	1	6	2	5	4	1	2
2 Teacher's knowledge of the diagnostic-prescriptive approach to reading has increased.	2	5		3	8		
3 Student achievement in reading has increased.	3	8	4	2	9	9	2
4 Teachers are more satisfied with the reading program.	4	10	7	8	9	9	10
5 The reading program is more successful.	5	7		1			
6 Teacher's use of the school system's curriculam guide has increased	6	3	1		6		6
7 Use of the system's curricular guide as a source of reading program goals increased.	7	4	3	7	7	2	4

TABLE 25
(Continued)

IMPACT STATEMENTS	READING COMMITTEE	RANK						
		SCHOOL NUMBER						
		320	334	336	360	368	411	
8 Students are more satisfied with the program.	8			9	9			
9 Principals are more satisfied with the reading program.	9		9	10				
10 The record keeping of reading skills is more effective.	10	9	6	5	2	1	3	

* W = 3.53   p < .017

significant at the .017 level. Therefore, the directional hypothesis
was retained and the null hypothesis was rejected.

Although there was overlap among the lists of impact statements,
the committee added the additional criterion that an item must appear
on all seven lists to be considered an undisputed impact of the READ
curriculum. Five statements were identified as appearing on all
seven lists:

1. Teacher's knowledge of the school system's curricular
   guide has increased.

2. The record keeping of reading skills is more effective.

3. Use of the system's curricular guide as a source of
   reading program goals has increased.

4. Student achievement in reading has increased.

5. Teachers are more satisfied with the reading program.

These five statements were considered validated effects of the
READ curriculum.

The impact statements generated from the metaevaluation study were
then compared with the findings related to hypothesis 1, 2, and 3.

The most direct comparison was with hypothesis 3, student achievement.
Significant differences were found in the mean reading achievement of
third and sixth graders in vocabulary, comprehension, and in total
reading using an analysis of covariance. The inference that these
effects were due to the READ curriculum is supported by impact statement
#4, student achievement in reading has increased. Likewise, for hypothesis
1, positive effects were found for Teachers' and Principals' Knowledge
and Practices. Impact statement #1 attributed knowledge of the system's
curricular guide to the READ curriculum. This was therefore considered
a validated effect of the READ curriculum upon principals.

Practices relative to the uses of the curricular guide could also be inferred for organizing and instructing students in reading. Since the guide contains the philosophy of the reading program and the scope and sequences of reading skills, it may therefore influence what teachers teach and the methods of teaching. In addition, adoption of the lead basal and co-basal textbooks, attributable to the READ curricular, may have influenced the methods of teaching since the response of teachers did indicate that a basal reading approach had replaced the phonic approach as the most frequently used method of instruction. The record keeping of reading skills which was also attributed to the READ curriculum was inferred to have been a factor that had influenced teacher evaluation practices. The record-keeping system utilized a checklist whereby teachers indicated reading level and mastery of skills. This information would be useful for evaluating a student's growth in reading as well as for aiding a teacher in organizing for reading instruction. Parents' satisfaction with the reading program was not validated. Conversely, teachers' satisfaction was an impact statement, although the cluster of pretest and posttest responses did not differ significantly between 1979 and 1983 when the sign test was applied. Table 26 shows the summary of findings from the pretest/posttest study and the impact statements from the metaevaluation study.

TABLE 26

Summary of Findings from the Pretest Posttest Study and
Impact Statements from the
Metaevaluation Study

FINDINGS FROM PRETEST/POSTTEST STUDY	IMPACT STATEMENTS FROM METAEVALUATION STUDY
Hypothesis 1	**Hypothesis 4**
1. Teachers' and principals' knowledge and skills	Teachers' knowledge of the curricular guide has increased.
2. Instructing students in reading	Use of the system's curricular guide as a source of reading program goals has increased.
3. Methods of teaching	Use of the system's curriculum guide as a source of reading program goals has increased.
4. Teacher evaluation practices	The record keeping of reading skills is more effective
5. Organizing students for reading instruction	The record keeping of reading skills is more effective.
Hypothesis 2	
6. More parents expressed satisfaction with the reading program. In general more principals, teachers, parents and students expressed more satisfaction with the reading program.	More teachers expressed satisfaction with the reading program.
Hypothesis 3	
7. Student achievement in reading has increased	Student achievement in reading has increased

In summary the findings related to the four major hypothesis were as follows:

Hypothesis 1. There was a difference in the description of the reading program after the implementation of the READ curriculum. These differences were organizing for instruction, instructing students, methods of teaching, teacher evaluation practices, and teachers' and principals' knowledge and practices.

Hypothesis 2. There was a significant difference in the overall satisfaction of principals, teachers, parents, and students with the reading program.

Hypothesis 3. There was a significant difference in the reading achievement of third and sixth grade students after the implementation of the READ curriculum. Significant differences ($p \leq .05$) were found for both cohorts in vocabulary, comprehension, and total reading.

Hypothesis 4. Independent groups of school staff did agree on changes in the reading program that could be attributed to the READ curriculum.

CHAPTER V

SUMMARY, CONCLUSIONS, AND RECOMMENDATIONS

Summary

The purpose of the study was to determine the impact of curricular intervention on the reading program of the Guilford County School System. Baseline data about the program were collected in 1979. The READ curriculum was implemented in 1981 in an attempt to carry out the recommendations stemming from the 1979 study. The READ curriculum consisted of a curricular guide, basal reader and co-basal reading series, checklist of reading skills, periodic and cumulative tests, and a reporting system for parents. This study sought to determine whether the perceptions of principals, teachers, parents, and students regarding the description of the reading program and satisfaction with the program were the same after the implementation of the READ curriculum as they were in 1979. In addition, the study sought to determine whether the reading achievement of the students had increased. In order to strengthen the internal and external validity of the study, a metaevaluation technique was employed. Multiple, independent replications of two rounds of impact questionnaires sought to determine whether multiple audiences would "construct" the same list of impact statements regarding the READ curriculum.

The sample consisted of students, parents, teachers, and principals in a North Carolina school system. Thirty-one principals, 605 teachers, 754 parents, and 1,149 students responded to the pretest surveys while 44 principals, 270 teachers, 918 parents, and 1,131 students responded to the posttest surveys. In addition, six teachers from six randomly selected schools participated in the metaevaluation study along with the district level Reading Curriculum Committee. The two student populations selected to determine reading achievement were the third and sixth grade students from the system's elementary and middle grade schools.

Reading surveys were developed locally to gather the perceptual data related to the reading program. The Principal and Teacher Surveys were designed to provide information about the description of the reading program with respect to organizing for instruction, instructing the students, goals and objectives, planning for instruction, methods of teaching and evaluation practices. The Parent, Student, Principal, and Teacher Surveys all contained questions regarding satisfaction with the reading program. Student achievement data were obtained from administration of the California Achievement Test.

The two major issues were addressed by the study: How can a local school district determine the effects of curriculum intervention in reading? And how can it be validated that these effects were due to the intervention and not to other factors?

A review of the literature related to program evaluation and to the evaluation of reading programs was conducted in order to gain information regarding the first issue. Selected literature related to program evaluation revealed that most of the well-known evaluation models could be roughly classified into three major approaches: The achievement-of-desired-outcomes, the assessment-of-merit, and the decision-making models. The Tyler model was reviewed as an example of the achievement-of-desired-outcomes model. The major disadvantage to using this model for determining the impact of the READ curriculum was that the outcomes from this approach were limited to the student's achievement of objectives. The works of Scriven and of Stake were reviewed as examples of assessment of merit models. These models took a broader view of the reading program including but not limiting the study to such things as materials, methods of instruction, and teacher knowledge. The major disadvantage of the assessment-of-merit models may be the lack of objectivity of the judgments stemming from some such evaluations. The CIPP model, system assessment, and the discrepancy models were reviewed as decision-making models of program evaluation. These models were presented as total approaches which serve program planning, program implementation, and program evaluation. It was noted that these three major evaluation models were not mutually exclusive and that advantages could be gained from using various components of the models for specific problem needs.

A review of the literature related to metaevaluation was conducted in order to address the second issue of the study, validation of the causal inference that changes in the program were due to the READ curriculum. Metaevaluation was defined by Cook and Gruder (1968), as

the evaluation of summative evaluation studies where the data are collected directly from program participants within a systematic design framework. The purpose of a metaevaluation was to improve the technical quality of an evaluation. Cook and Gruder described seven major models of metaevaluation including multiple independent replications, the model selected to address the second issue tapped by this study.

Four major hypotheses were tested for the study: hypothesis 1 related to describing the program differently, hypothesis 2 regarding satisfaction or more positive attitudes toward the reading program, hypothesis 3 concerning higher reading achievement, and hypothesis 4 related to the validation of impact statements about the reading program. The sign test was applied to pretest and posttest survey data to determine significantly different responses to questions describing the reading program. Significantly different responses were found for teachers on the items organizing students for reading instruction, instructing students in reading, methods of teaching, teacher evaluation practices, and teacher's and principal's knowledge and practice. For principals, the only significantly different responses were related to knowledge and practice. For these areas of the program, the directional hypothesis was supported. The null hypothesis was retained for other clusters of items on the surveys including goals and objectives and planning for instruction. The sign test was also used to determine the satisfaction of students, parents, teachers and principals with respect to the reading program. The only significantly different set of responses between 1979 and 1983 related to satisfaction with the program was from the parents. For parents, hypothesis #2 was supported, but

for principals, teachers, and students the directional hypothesis was rejected. For hypothesis #3, mean reading achievement scores attained in 1982 were compared to those attained by students in 1979. The analysis of covariance yielded F ratios of 47.74 for third grade vocabulary, 44.45 for third grade comprehension and 39.87 for third grade total reading. These F ratios were all significantly different than the hypothesized value of .05. Similar results were identified for sixth grade vocabulary, comprehension and total reading. Reading achievement scores were higher for students in grades three and six. Therefore, hypothesis #3 stating that there would be a difference in reading achievement significant at the .05 level was supported.

Hypothesis #4, stating that there would be no difference in seven independently derived lists of impact statements related to the READ curriculum was tested by using Kendall's Coefficient of Concordance. Seven lists of impact statements were generated by six randomly selected schools and by the Reading Curriculum Committee. With W=3.53, p $<$ .017, there was more agreement among the lists than was expected by chance. Therefore, hypothesis #4 was retained. Five impact statements were ranked among the top of each list and were therefore considered to be impacts of the READ curriculum:

1. Teachers' knowledge of the school system's curricular guide has increased.

2. The record keeping of reading skills is more effective.

3. Use of the system's curricular guide as a source of reading program goals has increased.

4. Student achievement in reading has increased.

5. Teachers are more satisfied with the reading program.

<u>Conclusions</u>

The findings of the study support the following conclusions:

1. Although decision-making models dominate the literature
   related to the evaluation of reading programs, the
   assessment-of-merit models and the achievement of
   desired-outcomes models are used as well. Both the
   purpose of the evaluation and the kind of information
   needed influence the selection of the evaluation model
   or combination of models that are used for the evaluation.

2. Of the covariates used in the study, mean school IQ was a
   better predictor of reading achievement than were percentage
   of minority students, percentage of students receiving help
   in Chapter I reading programs, and percentage of parents who
   had not completed high school.

3. Evaluations at the local level can be used to determine
   the effects of curricular intervention. The following effects of
   the READ curriculum   were determined by the pretest-
   posttest study and   were validated by the metaevaluation:
   teachers'   knowledge of the school system's curricular
   guide increased; the record keeping of reading skills was
   more effective; use of the system's curricular guide as
   a source of reading program goals had increased; student
   achievement in reading had increased; and teachers were
   more satisfied with the program.

4. Metaevaluation techniques are useful methods for validating
   the findings from field study evaluations where quasi-or
   true experimental designs cannot be used.

Recommendations for Future Research

A great deal has been written about metaevaluation techniques, but few school districts have used these techniques although their expressed purpose is to improve the technical quality of evaluations. This study has demonstrated that the validity and reliability of a school system evaluation was enhanced by using a metaevaluation technique : independent multiple replications. In general, a greater emphasis on the validation of program effects for local school districts is needed. Metaevaluation techniques have the potential for meeting this need.

In future studies that employ multiple-independent replications as a means of validating program effects, it is recommended that parents, principals, and students be included in the validation process. One of the findings from the pretest-posttest study was that parents expressed more satisfaction with the reading program. Only teachers were used in the validation study and it may have been difficult for them to project how parents viewed the reading program. If parents had been included in this part of the study, more definitive statements could be made about parental satisfaction with the reading program. Rauch (1983) supported the use of parents on the evaluation team. Including parents, principals, and students in the validation process would broaden the perspective of the validation process.

There are other strategies that would improve the multiple independent replications model. Consecutive independent replications hold promise for future studies of this kind. The evaluator may discover strategies during one replication which may be tested and applied during subsequent replications, thus improving the technical

quality of the study while it is in progress. In the present study parents could have been surveyed during a consecutive replication if all of the replications had not been conducted simultaneously. This would have added to validation or lack of validation of parental satisfaction with the reading program.

The combined use of perceptual data and data collected by trained observers would likewise enhance the objectivity of studies of this kind. This technique would be especially effective if the perceptual data and the observational data produce the same findings. These are techniques to be considered for future research.

BIBLIOGRAPHY

Alkin, M. C. Evaluation theory development. In C. H. Weiss (Ed.),
   Evaluating action programs: Reading in social action and
   education. Boston: Allyn & Bacon, Inc., 1972.

Altschuld, J. W., & C. V. Hines. Factors affecting the validity of
   field tests in education. Educational Evaluation and Policy
   Analysis, 1982, 4(3), 331-339.

Beauchamp, G. A. Curriculum theory. Wilmette, Illinois: The Kagg
   Press, 1975.

Beck, M. D. An adaptation of the standards for evaluations for the
   interpretation of reading program evaluations (Final Report).
   Bethesda, Md.: ERIC Document Reproduction Service, ED214110, 1981.

Benjamin, P. W. Cooperative development of planning and evaluation
   skills. Educational Leadership, 1978, 35(5), 401-403.

Borich, G. D., & R. P. Jamelka. Definitions of program evaluation
   and their relation to instructional design. Educational Technology,
   21(8), 31-38.

California Achievement Tests (Technical Bulletin, No. 1). Monterey,
   California: CTB McGraw-Hill, 1979.

Campbell, D. T., & J. C. Stanley. Experimental and quasi-experimental
   designs for research on teaching. In N. L. Gage (Ed.), Handbook of
   Research on Teaching. Chicago: Rand McNally, 1963.

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

Cook, T. D., & C. L. Gruder. Metaevaluation research. Evaluation
   Quarterly, 1978, 2(1), 5-51.

Cronbach, L. J. Course improvement through evaluation. In B. J. Worthen
   and J. R. Sanders (Eds.), Educational evaluation theory and practice.
   Belmont, California: Wadsworth, 1973.

Daniel, W. W. Applied nonparametric statistics. Boston: Houghton-
   Mifflin, 1978.

District of Columbia Public Schools. Evaluation of outcomes, 1976-1977:
   An evaluation system report on reading programs and reading
   achievement (Part 1A Summary). Bethesda, Maryland: ERIC Document
   Reproduction Service, ED169491, 1979.

Dyer, H. S., Linn, R. L., & M. J. Patton. A comparison of four
methods of obtaining discrepancy measures based on observed
and predicted school system means on achievement tests. American
Educational Research Journal, 1969, 6, 591-605.

Elashoff, J. D. Analysis of covariance: A delicate instrument.
American Educational Research Journal, 1969, 6(3), 385.

Ellis, A. B. Success and failure. Watertown, Massachusetts:
Education Research Corporation, 1975.

Garberina, T. J. A systems approach for generating change in the
reading program. Bethesda, Maryland: ERIC Document Reproduction
Service. ED124927, 1976.

Goodrich, R. L., & R. G. St. Pierre. Opportunities for studying later
effects of follow through (Report prepared for the U. S. Office of
Education under contract no. 300-78-0443). Cambridge, Mass.: Abt
Associates, 1979.

Guba, E. G. The failure of educational evaluation. In C. H. Weiss
(Ed.), Evaluating action programs: Reading in social action and
education. Boston: Allyn & Bacon, Inc., 1972.

Guilford County School System. READ curriculum guide. Greensboro,
North Carolina: Author, 1979.

Hammond, R. L. Accountability through context evaluation of reading
performance. Journal of Research and Development in Education,
1971, 5(1), 31-36.

Haney, W. A technical history of the national follow through evaluation
(Report to the U. S. Office of Education pursuant to contract no.
OEC-0-74-0394). Cambridge, Mass.: Huron Institute, 1977.

Haney, W. The follow through planned variation experiment: Statistical
conclusion validity? Paper presented at the annual meeting of the
American Educational Research Association, Toronto, Ontario, 1978.

Haney, W., & N. Pennington. Reanalysis of follow through parent and
teacher data from spring 1975 (Report to the U. S. Office of
Education pursuant to contract no. OEC-74-0394). Cambridge, Mass.:
Huron Institute, 1978.

Huebner, M. Innovations in reading: A look at evaluative procedures.
In J. A. Figurel (Ed.), Vistas in reading. Newark, Delaware:
International Reading Association, 1967.

Hock, M. D., Sellers, H. W., Blatt, S., & O. Gault. A report on the
Saginaw OSUEC Collaborative Project in Evaluation, August, 1968 to
June, 1969. In D. L. Stufflebeam, W. J. Foley, W. J. Gephant,
E. G. Guba, R. L. Hammond, H. O. Merriman, & M. M. Provus, (Eds.),
Education evaluation and decision making. Itasca, Illinois:
F. E. Peacock, 1971.

House, E. R. Assumptions underlying evaluation models. _Educational Researcher_, 1978, _7_(3), 4-12.

House, E. R. No simple answer: Critique of the follow through evaluation. _Harvard Educational Review_, 1978, _48_(2), 128-160.

Hull, C. H., & N. H. Nie. _SPSS statistical package for the social sciences_ (Version M, Release 9.1). New York: McGraw Hill, 1982.

Joint Committee on Standards for Educational Evaluation. _Standards for evaluations of educational programs, projects, and materials_. New York: McGraw-Hill, 1981.

Kennedy, M. M. Generalizations of findings from single-case studies. Paper presented at the annual meeting of the AERA, Toronto, 1978.

Mason, P. A. _Reading in reading: A program evaluation_ in Reading Public Schools, Massachusetts. Bethesda, Md.: ERIC Document Reproduction Service: ED209281, 1981.

McCormick, W. J. _Developing and implementing a process for identifying needs in a reading program as a basis for a plan of corrective action_. Bethesda, Md.: ERIC Document Reproduction Service, ED127565, 1976.

McGuire, M. L. _Assessment of reading programs_. Bethesda, Md.: ERIC Document Reproduction Service, ED027150, 1968.

Millman, J. A checklist procedure. In N. L. Smith (Ed.), _New techniques for evaluation_. Beverly Hills: Sage, 1981.

Mills, P., & L. Crawford. _Fort Gay-Thompson evaluation project_. Bethesda, Md.: ERIC Document Reproduction Service, ED15118, 1973.

Pedhazur, E. J. _Multiple regression in behavioral research explanation and prediction_. New York: Holt, Rinehart and Winston, 1982.

Provus, M. M. Discrepancy evaluation model 1969. In D. L. Stufflebeam, W. J. Foley, W. J. Gephart, E. G. Guba, R. L. Hammond, H. O. Merriman, & M. M. Provus (Eds.), _Educational evaluation and decision making_. Itasca, Illinois: F. E. Peacock, 1971.

Provus, M. M. Evaluation of ongoing programs in the public school system. In B. R. Worthen & J. R. Sanders (Eds.), _Educational evaluation: Theory and practice_. Belmont, California: Wadsworth Publishing Co., 1973.

Ramsey, J. G. _Supplementary reading program for handicapped children 1975-76: Evaluation report_. Bethesda, Md.: ERIC Document Reproduction Service, ED136485, 1976.

Rauch, S. The administrator and the reading program lessons learned by an evaluator. _Reading World_, 1983, _22_(3), 255-257.

Rogers, F. A. Curriculum research and evaluation. In F. W. English
   (Ed.), _Fundamental curriculum decisions_. Alexandria, Va.: Association
   for Supervision and Curriculum Development, 1983.

Scriven, M. _The methodology of evaluation_. (AERA Monograph Series on
   Curriculum Evaluation, No. 1). Chicago: Rand McNally, 1967.

Scriven, M. The methodology of evaluation. In C. H. Weiss (Ed.),
   _Evaluating action programs: Readings in social action and
   education_. Boston: Allyn & Bacon, 1972.

Scriven, M. Evaluation perspectives and procedures. In W. J. Popham
   (Ed.), _Evaluation in education: Current applications_. Berkeley:
   McCutchan, 1974.

Scriven, M. Evaluation bias and its control. In G. V. Glass (Ed.),
   Studies review annual (Vol. 1). Beverly Hills, California: Sage
   Publications, 1976.

Smith, N. L. _New techniques for evaluation_. Beverly Hills: Sage
   1981.

Stake, R. E. The countenance of educational evaluation. _Teachers
   College Record_, 1967, _68_, 523-540.

Stake, R. E. _Case-study method in social inquiry_. Champaign:
   University of Illinois, 1976.

St. Pierre, R. G. Correcting covariates for unreliability: Does it
   lead to differences in an evaluator's conclusions? _Evaluation
   Quarterly_, 1978, _2_(1), 401-420.

St. Pierre, R. G. Follow through: A case in metaevaluation research.
   Educational Evaluation and Policy Analysis, 1982, _4_(1).

Stufflebeam, D. L., Foley, W. J., Gephart, W. J., Guba, E. G., Hammond,
   R. L., Merriman, H. O., & M. M. Provus. _Educational evaluation and
   decision making_. Itasca: F. E. Peacock Publishers, Inc., 1971.

Stufflebeam, D. L. The relevance of the CIPP evaluation model for
   educational accountability. _Journal of Research and Development
   in Education_, 1971, _5_(1), 19-25.

Stufflebeam, D. L. Metaevaluation: Concepts, standards, and uses.
   In R. A. Berk (Ed.), _Educational evaluation methodology: The state
   of the art_. Baltimore: The Johns Hopkins University Press, 1981.

Taylor, P. A., & T. O. Maguire. A theoretical evaluation model. Cited
   by R. E. Stake in C. H. Weiss (Ed.), _Evaluating action programs:
   Reading in social action and education_. Boston: Allyn & Bacon, Inc.,
   1972.

Tyler, R. W. Basic principles of curriculum and instruction. Chicago:
    The University of Chicago Press, 1949.

U. S. Office of Education. Right to read assessment and planning
    handbook. Washington, D.C.: Government Printing Office, 1974.

Wallen, C. J., & S. R. Wisely. Analysis of elementary reading programs.
    Bethesda, Md.: ERIC Document Reproduction Service, ED046623,
    1970.

Worthen, B. R., & J. F. Sanders. Educational evaluation: Theory and
    practice. Belmont, California: Wadsworth, 1973.

APPENDIX A

PRINCIPAL SURVEY IN READING

COVER LETTER

PERCENTAGE OF RESPONSES 1979 and 1982

M E M O

TO: Principals of Grades K-12

FROM: Juanita Johnson
        Instructional Supervisor of Reading

DATE: January 13, 1983

RE: READING CURRICULUM EVALUATION and IMPROVEMENT

I would like to solicit your participation and support in the
continuous evaluation and improvement of our reading program. Your
participation is requested in the following ways:

1. Select a representative to serve on the Systemwide Reading
    Curriculum Committee. This committee will have as its task the
    continuous evaluation and improvement of the reading program.
    This committee will meet on Wednesday, February 2, 1983, at
    2:45 p.m. in the Board Room of the Administrative Offices. The
    meeting will last until 4:15 p.m.

2. Select a school based reading curriculum committee (if you do
    not already have one) which may be chaired by you or the person
    whom you select to serve on the systemwide committee. This
    committee will be responsible for evaluating the reading
    program as it is operating within the school and for making
    recommendations for improvement.

3. Encourage your faculty to cooperate with the school committee
    as well as the systemwide committee.

4. Take the time to answer the questions on the Principal's Survey
    in Reading which you will be sent and return it as soon as
    possible.

EDUCATIONAL EXCELLENCE

MEMO
Principals of Grades K-12
January 13, 1983
Page 2

This procedure for evaluating and making recommendations for
improving our reading program is suggested as a way of relieving
you, the principal, of some of the time-consuming tasks such as the
distribution and collection of surveys. It is also seen as a means
to actively involve more teachers in the curricular improvement
process.

Thank you for your help and your continued support. It is only
through your help that our reading program will be improved.

C. Howard Cross, Assistant Superintendent/
  Elementary Education

T. G. Madison, Assistant Superintendent/
  Secondary Education

PRINCIPAL   SURVEY IN READING

DIRECTIONS: Choose the appropriate answer for the following statements and mark the
             correct box.

A. I have done this often.   1983 > 1979 = +
B. I have done this occasionally.   1979 > 1983 = -
C. I have not done this.   No signs are attached to ties
D. I have not done this but think I should.   *Significant at the .05 level
E. Not applicable.

I. Program Organization	Percent of Responses 1979 n=32					Percent of Responses 1983 n=44					DIRECTION OF DIFFERENCE
	A	B	C	D	E	A	B	C	D	E	
A. In Organizing Students for Reading Instruction...My Teachers											
1. Divide students into more than one group on the basis of reading ability (i.e., ability grouping).	78	13	3	3	3	85	10	3	3		+
2. Group students on the basis of specific reading needs (i.e., special needs grouping).	78	19			3	51	41	8			−
3. Base classroom reading instruction on the idea of whole-class grouping (i.e., one group).	9	19	59	3	9	16	27	47	5	5	+
4. Instruct students individually in reading rather than in groups (i.e., individualized reading instruction).	28	53	13			6	27	54	9	9	0
5. Group students for reading instruction on the basis of commonly-shared interests (i.e., interest grouping).	13	38	28	9	6	7	54	32	9		+
6. Use the results of standardized reading achievement tests as measured by the State Annual Testing Program.	50	41	6	3		50	32	11	7		−
7. Move students from one group to another (i.e., flexible grouping) as needs vary.	75	22			3	73	23	5			−
8. Group students on the basis of the group they were in last year.	6	22	66	3	3	14	36	36	11	2	+
9. Organize my reading program with the help of reading specialists.	34	31	22	9	3	10	40	25	15	10	−

DIRECTIONS: Choose the appropriate answer for the following statements and mark the correct box.

A. I have done this often.
B. I have done this occasionally.
C. I have not done this.
D. I have not done this but think I should.
E. Not applicable.

1983 > 1979 = +
1979 > 1983 = —
No signs are attached to ties
*Significant at the .05 level

	Percent of Responses 1979 (n=32)					Percent of Responses 1983 (n=44)					DIRECTION OF DIFFERENCE
	A	B	C	D	E	A	B	C	D	E	
10. Use parents, paraprofessionals and/or community resources.	44	41	13	6		25	57	9	9		0
11. Organize my reading program with the help of colleagues.	50	34	16			71	21	3	3	3	+

P = .5001 N.S

B. Concerning Goals and Objectives...
My Teachers

	1979 A	B	C	D	E	1983 A	B	C	D	E	Dir.
12. Use the school system's curriculum guide in reading as the source of reading program goals.	50	28	6	6	6	54	32	7	5	2	+
13. List general goals for the reading program based on their assessment of the students' reading strengths and reading needs.	59	25	9	6		57	32	5	7		+
14. Maintain a record keeping system to keep track of individual progress toward specific objectives.	63	25	6	3	3	72	19	5	5		+
15. Use the basal reader as the source of their reading program goals.	59	22	9		9	80	16		2	2	+

P = .0625

C. In Planning for Skills Instruction...
My Teachers

	1979 A	B	C	D	E	1983 A	B	C	D	E	Dir.
16. Organize classroom reading instruction on the basis of skill levels represented in their classes.	75	22		3		84	11	2	2		—
17. Establish a sequence of reading skills based on assessment of student reading needs.	72	19	6	3		73	24		2		+
18. Use the basal reading series to determine the sequence in which reading skills are taught.	34	41	16	3	3	63	31	3		3	+

DIRECTIONS: Choose the appropriate answer for the following statements and mark the correct box.

A. I have done this often.
B. I have done this occasionally.
C. I have not done this.
D. I have not done this but think I should.
E. Not applicable.

1983 > 1979 = +
1979 > 1983 = -
No signs are attached to ties
*Significant at the .05 level

| | Percent of Responses | | | | | | | | | | DIRECTION OF DIFFERENCE |
| | 1979 n=32 | | | | | 1983 n=44 | | | | | |
	A	B	C	D	E	A	B	C	D	E	
19. Select practice activities that match their instructional objectives.	88	13				93	5		2		+
20. Organize their classroom with learning centers.	47	44	6		3	42	47	5	5	2	-

P = .4999

D. **When Instructing Students in Reading**
    ...My Teachers

	A	B	C	D	E	A	B	C	D	E	
21. Teach reading through non-basal materials.	56	44				42	49	7		2	-
22. Assign workbook pages as practice activities which match instructional objectives.	72	19	6		3	70	26		2	2	+
23. Use a file of workbook pages and exercises classified by skill and level.	50	34	9	3	3	58	37		5		+
24. Use workbooks as the major guide to introducing reading skills.	9	41	47		3	21	48	26	5		+
25. Have teacher-pupil planning sessions characterized by considerable give-and-take.	19	44	19	13	3	7	49	30	14		-
26. Prepare a directory of commercial reading materials available within my school.	25	9	31	34		28	26	26	19	2	+
27. Use the State's Guide for Evaluation of Materials.	16	22	34	19	9	30	35	26	7	2	+
28. Code reading materials to reading objectives.	41	25	19	13	3	36	33	21	7	2	+
29. Use audio-visual materials.	81	19				73	23	2	2		-

P = .2539

DIRECTIONS: Choose the appropriate answer for the following statements and mark the correct box.

A. I have done this often.
B. I have done this occasionally.
C. I have not done this.
D. I have not done this but think I should.
E. Not applicable.

1983 > 1979 = +
1979 > 1983 = –
No signs are attached to ties
*Significant at the .05 level

II. Instructional Practices

A. Methods of Teaching...My Teachers

| | Percent of Responses | | | | | | | | | | DIRECTION OF DIFFERENCE |
| | 1979 n=32 | | | | | 1983 n=44 | | | | | |
	A	B	C	D	E	A	B	C	D	E	
30. Use the basal approach.	84	9	3		3	84	7		2	7	–
31. Use the language experience approach.	47	38	6	3	3	29	57	5	2	7	+
32. Use the phonics approach.	66	28	3			55	36	2		7	–
33. Use programmed instruction.	22	50	19		6	31	43	21		5	+
34. Use an individualized approach.	50	44	3	3		28	67	5			+
35. Use management systems. EXAMPLE: PRI.	50	22	16	6	3	31	45	14	2	7	+
36. Use an eclectic approach.	50	19	13		3	35	38	24	2		+
37. Other (list on number 73).											
38. Encourage a child to select topics he/she or a group may wish to read about.	47	47			3	52	43	2	2		+
39. Release a child from group work to do individual reading.	34	59				38	50	7	2	2	–

P = .2539

III. Teacher Evaluation Practice...My Teachers

	A	B	C	D	E	A	B	C	D	E	
40. Use records of independent reading done by each student.	41	41	9			36	55	9			+
41. Consider a child's ability to discuss what he/she has heard others read aloud.	56	34			3	48	52				+
42. Consider work in reading other than the basal program material.	78	22				58	35		2	5	–
43. Compare achievement to behavioral objectives by means of a criterion-referenced test (the objectives state the conditions under which a child should do something which teachers can observe, to a degree which the teachers have specified).	31	38	9	9	9	21	58	12	9		+

DIRECTIONS: Choose the appropriate answer for the following statements and mark the correct box.

A. I have done this often.
B. I have done this occasionally.
C. I have not done this.
D. I have not done this but think I should.
E. Not applicable.

1983 > 1979 = +
1979 > 1983 = −
No signs are attached to ties
*Significant at the .05 level

		Percent of Responses 1979 n=32					1983 n=44					DIRECTION OF DIFFERENCE
		A	B	C	D	E	A	B	C	D	E	
44.	Use results of an informal reading inventory.	44	41	6	3		43	50	8			+
45.	Consider to what degree a student's textbooks or required reading are matched to his reading level.	81	16	3			60	40				+
46.	The child and his/her teacher make a "performance contract" and the teacher assesses the child's progress in completing this contract.	16	53	19			6	10	59	26	3	0

P = .1094

IV. Principal's Knowledge

Use the following choices for questions 47-65.

A. Much, B. Some, C. Little, D. None

		A	B	C	D	E	A	B	C	D	E	
47.	I have training in teaching reading.	25	36	25	13		44	39	10	5	2	+
48.	I have knowledge of the diagnostic-prescriptive approach to reading.	38	53	6	3		49	47	2	2		+
49.	We use a diagnostic prescriptive approach to teaching reading.	38	41	19	3		26	55	8	8	3	+
50.	I have knowledge of motivation techniques.	41	53	6			60	40				+
51.	We use these techniques.	44	44	13			43	48	7		2	+
52.	I have knowledge of our school system's curriculum guide.	66	28			3	3	67	23	7	2	−
53.	We use this guide.	38	47	3	6	6	6	64	26	7	2	+
54.	I have knowledge of management system techniques.	38	41	13	9		51	39	7	2		+
55.	We use a management system.	19	25	31	19	6	39	39	20	2		+
56.	I have knowledge of standardized norm-referenced tests.	38	56	6			46	44	10			−
57.	We use the results of standardized norm-referenced tests.	22	56	16	3	3	43	40	12		5	+

1983 > 1979 = +
1979 > 1983 = −
No signs are attached to ties
*Significant at the .05 level

	Percent of Responses 1979 n=32					Percent of Responses 1983 n=44					DIRECTION OF DIFFERENCE
	A	B	C	D	E	A	B	C	D	E	
58. I have knowledge of criterion-referenced tests.	31	50	16		3	42	51	7			+
59. We use the results of criterion-referenced tests.	25	38	34		3	33	48	12	2	5	+
60. I have knowledge of criteria for selecting materials.	28	53	16		3	41	51	7			+
61. We use criteria for selecting materials.	31	50	16		3	39	49	10		2	+
62. I have knowledge of the sources for reading materials.	38	59	3			48	48	2	2		−
63. I have knowledge of the issue of accountability and its implications.	53	47				55	38	7			−

*P = .0245

DIRECTIONS: Choose the appropriate answer for the following statements and mark the correct box.

A. Very
B. Fairly
C. Big Problems
D. No Response
E. Not Applicable

V. Principal's Evaluation of the Reading Program

	Percent of Responses 1979 n=32					Percent of Responses 1983 n=44					DIRECTION OF DIFFERENCE
	A	B	C	D	E	A	B	C	D	E	
64. How successful is your school's reading program on the whole?	63	38				58	40	2			. −
65. How satisfied are you with the reading skills and habits of the students in your classes?	38	59	3			48	50	2			+
66. How effective do you find the present reporting system for record keeping of reading skills as students move from grade to grade?	25	53	19	3		30	45	25			−

DIRECTIONS: Choose the appropriate answer for the following statements and mark the correct box.

A. Very
B. Fairly
C. Big Problems
D. No Response
E. Not Applicable

1983 > 1979 = +
1979 > 1983 = –
No signs are attached to ties
*Significant at the .05 level

| | Percent of Responses | | | | | | | | | | DIRECTION OF DIFFERENCE |
| | 1979 n=32 | | | | | 1983 n=44 | | | | | |
	A	B	C	D	E	A	B	C	D	E	
67. How effective do you find the present reporting system in reading for school to school transfer?	16	38	47			21	48	31			+
68. The reading program in your school is geared toward which group of students?											
A = Above Average	3					7					+
B = Average	28					25					–
C = Below Average	3					18					+
D = All	53					48					–
E = No Response	13										
69. What evidence do you consider in making judgements about the instructional levels of the children?											
A. Informal Reading Inventory	66					67					–
B. Individual conference	44					32					–
C. Publisher's tests											
D. Own tests	69					43					–
E. Books pupil chooses to read	38					25					–
F. Oral reading	56					57					+
G. Workbooks	22					27					+
H. Standardized Tests	63					61					–

P = .7728 NS

APPENDIX B

TEACHER SURVEY IN READING

COVER LETTER

AND

PERCENTAGE OF RESPONSES 1979 AND 1982

ROBERT M. BOGGS, SUPERINTENDENT

M E M O

TO: Principals and Teachers

FROM: T. G. Madison, Assistant Supt./Secondary Education
        Howard Cross, Assistant Supt /Elementary Education
        Juanita Johnson, Reading Supervisor

DATE: February 2, 1983

RE: Where Are We In Reading?
        Where Do We Want to Go?

The Guilford County Schools Administration has requested a follow-
up study of the evaluation of the reading program. The most
important part of the evaluation is the input of you, the
principals and teachers of the school system. The following
questionnaire has been designed by teachers, principals, and county
office personnel. It is hoped that your candid answers to the
questions will give a picture of our present program and a
direction for the future.

The answers given in the survey will be treated confidentially. To
insure confidentiality, we are asking that you not sign your names
nor designate your school.

We appreciate your cooperation and thank you for returning your
completed questionnaire to the Reading Supervisor by February 14,
1983.

Please keep in mind the following definition of reading as you
answer the questionnaire:

  Reading is the recognition and comprehension of written language.

EDUCATIONAL EXCELLENCE

GUILFORD COUNTY SCHOOL SYSTEM

TEACHER   SURVEY IN READING

DIRECTIONS: Choose the appropriate answer for the following statements and mark the
correct box.

A.   I have done this often.
B.   I have done this occasionally.
C.   I have not done this.
D.   I have not done this but think I should.
E.   Not applicable.

1983 > 1979 = +
1979 > 1983 = -
No signs are attached to ties
*Significant at the .05 level

I. Program Organization	Percent of Responses										DIRECTION OF DIFFERENCE
A. In Organizing Students for Reading Instruction...I	1979 n=641					1983 n=270					
	A	B	C	D	E	A	B	C	D	E	
1. Divide students into more than one group on the basis of reading ability (i.e., ability grouping).	81	10	3	0	6	81	11	3	1	5	+
2. Group students on the basis of specific reading needs (i.e., special needs grouping).	59	28	5	3	4	57	33	4	1	4	+
3. Base classroom reading instruction on the idea of whole-class grouping (i.e., one group).	14	22	51	1	10	20	29	40	1	7	+
4. Instruct students individually in reading rather than in groups (i.e., individualized reading instruction).	14	53	20	4	8	17	55	8	3	4	+
5. Group students for reading instruction on the basis of commonly-shared interests (i.e., interest grouping).	3	27	50	10	9	5	33	46	7	7	+
6. Use the results of standardized reading achievement tests as measured by the State Annual Testing Program.	27	31	22	3	14	31	38	15	4	12	+
7. Move students from one group to another (i.e., flexible grouping) as needs vary.	57	30	4	1	7	58	28	4	4	6	-
8. Group students on the basis of the group they were in last year.	7	20	40	1	15	16	28	39	1	14	+
9. Organize my reading program with the help of reading specialists.	6	21	40	8	17	8	30	44	8	11	+
10. Use parents, paraprofessionals and/or community resources.	20	40	23	9	7	25	46	21	4	3	+

DIRECTIONS: Choose the appropriate answer for the following statements and mark the correct box.

A. I have done this often.
B. I have done this occasionally.
C. I have not done this.
D. I have not done this but think I should.
E. Not applicable.

1983 > 1979 = +
1979 > 1983 = -
No signs are attached to ties
*Significant at the .05 level

	Percent of Responses 1979 n=641					Percent of Responses 1983 n=270					DIRECTION OF DIFFERENCE
	A	B	C	D	E	A	B	C	D	E	
11. Organize my reading program with the help of colleagues.	30	44	19	2	5	32	44	15	2	6	+

*P = .0059

B. Concerning Goals and Objectives...I

	A	B	C	D	E	A	B	C	D	E	
12. Use the school system's curriculum guide in reading as the source of reading program goals.	39	33	12	5	8	60	22	8	5	4	+
13. List general goals for the reading program based on the assessment of the students' reading strengths and reading needs.	62	25	3	4	4	51	38	5	1	4	+
14. Maintain a record keeping system to keep track of individual progress toward specific objectives.	52	28	8	9	3	62	23	8	4	3	+
15. Use the basal reader as the source of reading program goals.	56	19	11	1	13	64	20	7		8	+

P = .0625

C. In Planning for Skills Instruction...I

	A	B	C	D	E	A	B	C	D	E	
16. Organize classroom reading instruction on the basis of skill levels represented in the class.	73	18	3	1	4	74	19	3	1	3	+
17. Establish a sequence of reading skills based on assessment of student reading needs.	62	25	5	2	5	53	25	7	2	14	-
18. Use the basal reading series to determine the sequence in which reading skills are taught.	40	29	16	2	12	54	30	9		6	+
19. Select practice activities that match instructional objectives.	87	9	1	1	1	87	12	-	1	1	+
20. Organize the classroom with learning centers.	34	42	14	4	4	38	43	11	4	4	+

P = .1874

DIRECTIONS: Choose the appropriate answer for the following statements and mark the correct box.

A. I have done this often.
B. I have done this occasionally.
C. I have not done this.
D. I have not done this but think I should.
E. Not applicable.

1983 > 1979 = +
1979 > 1983 = −
No signs are attached to ties
*Significant at the .05 level

		Percent of Responses 1979 n=641					1983 n=270					DIRECTION OF DIFFERENCE
D.	When Instructing Students in Reading...I	A	B	C	D	E	A	B	C	D	E	
21.	Teach reading through non-basal materials.	35	48	11	2	4	38	49	9	1	3	+
22.	Assign workbook pages as practice activities which match instructional objectives.	57	28	6	1	8	67	22	4	1	5	+
23.	Use a file of workbook pages and exercises classified by skill and level.	47	27	12	6	6	48	27	15	4	4	+
24.	Use workbooks as the major guide to introducing reading skills.	13	33	42	1	9	17	34	34	2	9	+
25.	Have teacher-pupil planning sessions characterized by considerable give-and-take.	8	39	31	13	8	13	44	28	10	5	+
26.	Prepare a directory of commercial reading materials available within my school.	7	10	54	19	9	10	14	53	16	7	+
27.	Use the State's Guide for Evaluation of Materials.	5	22	48	13	9	12	28	44	8	6	+
28.	Code reading materials to reading objectives.	29	32	22	7	6	30	31	28	6	6	0
29.	Use audio-visual materials.	59	36	3	0	2	61	34	2	1	1	0

*P = .0076

II. Instructional Practices

A. Methods of Teaching...I

		A	B	C	D	E	A	B	C	D	E	
30.	Use the basal approach.	60	18	8	1	10	74	12	6	0	9	+
31.	Use the language experience approach.	32	51	9	1	5	35	48	10	1	6	0
32.	Use the phonics approach.	56	29	7	1	5	45	39	8	0	8	−

DIRECTIONS: Choose the appropriate answer for the following statements and mark the correct box.

A. I have done this often.   1983 > 1979 = +
B. I have done this occasionally.   1979 > 1983 = -
C. I have not done this.   No signs are attached to ties
D. I have not done this but think I should. *Significant at the .05 level
E. Not applicable.

		Percent of Responses 1979 n=641					Percent of Responses 1983 n=270				DIRECTION OF DIFFERENCE
		A	B	C	D	E	A	B	C	D	E	
33.	Use programmed instruction.	19	35	31	2	6	22	39	29	1	8	+
34.	Use an individualized approach.	29	51	9	3	4	30	55	10	0	5	+
35.	Use management systems. EXAMPLE: PRI.	27	18	37	4	9	26	38	24	0	11	+
36.	Use an eclectic approach.	32	17	28	1	10	34	25	29	0	12	+
37.	Other (list on number 73).											
38.	Encourage a child to select topics he/she or a group may wish to read about.	34	45	8	5	6	34	47	9	5	4	+
39.	Release a child from group work to do individual reading.	25	49	14	3	7	22	53	14	3	7	+

*P = .0351

III. Teacher's Evaluation Practices...

		A	B	C	D	E	A	B	C	D	E	
40.	Use records of independent reading done by each student.	32	36	17	6	7	31	44	15	4	7	+
41.	Consider a child's ability to discuss what he/she has heard others read aloud.	54	35	5	1	4	53	36	5	0	5	0
42.	Consider work in reading other than the basal program material.	67	24	2	0	6	59	33	3	0	6	+
43.	Compare achievement to behavioral objectives by means of a criterion-referenced test (the objectives state the conditions under which a child should do something which teachers can observe, to a degree which the teachers have specified).	21	32	27	6	8	22	44	19	7	8	+
44.	Use results of an informal reading inventory.	37	42	11	3	6	33	50	7	3	7	+

DIRECTIONS: Choose the appropriate answer for the following statements and mark the correct box.

A. I have done this often.
B. I have done this occasionally.
C. I have not done this.
D. I have not done this but think I should.
E. Not applicable.

1983 > 1979 = +
1979 > 1983 = −
No signs are attached to ties
*Significant at the .05 level

		Percent of Responses 1979 n=641					1983 n=270					DIRECTION OF DIFFERENCE
		A	B	C	D	E	A	B	C	D	E	
45.	Consider to what degree a student's textbooks or required reading are matched to his reading level.	69	18	3	1	8	57	31	3	1	7	+
46.	The child and his/her teacher make a "performance contract" and the teacher assesses the child's progress in completing this contract.	6	25	47	11	10	5	28	52	6	9	+

*P = .0156

IV. Teacher's Knowledge

Use the following choices for questions 47-63.
A. Much, B. Some, C. Little, D. None

		A	B	C	D	E	A	B	C	D	E	
47.	I have training in teaching reading.	35	48	10	5	2	51	31	9	6	2	−
48.	I have knowledge of the diagnostic-prescriptive approach to reading.	30	50	13	6	1	45	41	6	6	2	+
49.	I use a diagnostic prescriptive approach to teaching reading.	25	41	19	13	3	30	44	12	10	5	+
50.	I have knowledge of motivation techniques.	34	55	8	2	2	54	37	5	2	3	+
51.	I use these techniques.	35	51	9	3	2	50	36	8	2	3	0
52.	I have knowledge of our school system's curriculum guide.	40	42	11	5	2	66	24	4	2	3	+
53.	I use this guide.	28	42	16	11	3	53	34	6	2	5	+
54.	I have knowledge of management system techniques.	21	34	24	16	5	38	36	10	11	5	+
55.	I use a management system.	17	28	24	26	5	30	35	16	13	5	+

1983 > 1979 = +
1979 > 1983 = −
No signs are attached to ties
*Significant at the .05 level

142

		Percent of Responses									DIRECTION OF DIFFERENCE
		1979 n=641					1983 n=270					
		A	B	C	D	E	A	B	C	D	E	
56.	I have knowledge of standardized norm-referenced tests.	20	49	22	7	2	34	44	11	6	5	+
57.	I use the results of standardized norm-referenced tests.	8	40	30	19	3	23	43	19	7	8	+
58.	I have knowledge of criterion-referenced tests.	17	41	27	13	3	33	39	13	10	5	+
59.	I use the results of criterion-referenced tests.	9	36	30	22	3	17	44	19	13	7	+
60.	I have knowledge of criteria for selecting materials.	27	47	17	6	3	37	40	14	4	4	+
61.	I use criteria for selecting materials.	27	46	16	7	3	34	42	15	4	4	+
62.	I have knowledge of the sources for reading materials.	37	50	8	2	3	52	37	7	2	3	+
63.	I have knowledge of the issue of accountability and its implications.	33	48	11	5	3	56	31	8	2	3	+

*P = .0002

DIRECTIONS: Choose the appropriate answer for the following statements and mark the correct box.

    A. Very
    B. Fairly
    C. Big Problems
    D. No Response
    E. Not applicable.

V. Teacher's Evaluation of the Reading Program

		A	B	C	D	E	A	B	C	D	E	
64.	How successful is your school's reading program on the whole?	53	39	3	5		67	31	3			+
65.	How satisfied are you with the reading skills and habits of the students in your classes?	34	53	9	4		44	50	6			+
66.	How effective do you find the the present reporting system for record keeping of reading skills as students move from grade to grade?	17	56	21	5		24	55	20			+
67.	How effective do you find the present reporting system in reading for school to school transfer?	7	52	35	6		12	60	24			+

DIRECTIONS: Choose the appropriate answer for the following statements and mark the correct box.

A. Very
B. Fairly
C. Big Problems
D. No Response
E. Not applicable.

1983 > 1979 = +
1979 > 1983 = –
No signs are attached to ties
*Significant at the .05 level

	Percent of Responses										DIRECTION OF DIFFERENCE
	1979 n=641					1983 n=270					
	A	B	C	D	E	A	B	C	D	E	
68. The reading program in your school is geared toward which group of students?											
A = Above Average	4					3					–
B = Average	40					37					–
C = Below Average	9					8					–
D = All	43					51					+
E = No Response	5					0					–
69. What evidence do you consider in making judgements about the instructional levels of the children?											
A. Informal Reading Inventory	67					47					–
B. Individual Conference	46					33					–
C. Publisher's Tests	26					31					+
D. Own Tests	72					42					–
E. Books Pupil Chooses to Read	23					19					–
F. Oral Reading	76					57					–
G. Workbooks	26					25					–
H. Standardized Tests	44					43					–

P = .9617

APPENDIX C

PARENT SURVEY IN READING

COVER LETTER

PERCENTAGE OF RESPONSES 1979 AND 1982

Dear Parents,

The Guilford County School System has requested a follow-up study of the evaluation of the reading program. An important part of the evaluation is the input of the parents of Guilford County students. You have been chosen by a random sampling of Guilford County parents to participate in this study.

The following questionnaire has been designed for this purpose by teachers, principals, and county office personnel. It is hoped that your candid answers will give a picture of our present program and a direction for the future.

The answers given in the survey will be treated confidentially. To insure this confidentiality and for your convenience, an envelope has been provided. We would appreciate your returning this survey by February 14, 1983.

Thank you,

C. Howard Cross
Assistant Superintendent
  for Elementary Education

T. G. Madison
Assistant Superintendent
  for Secondary Education

EDUCATIONAL EXCELLENCE

GUILFORD COUNTY SCHOOL SYSTEM

PARENT SURVEY IN READING

DIRECTIONS: Choose the appropriate answer for the following statements and mark the correct box.

A. Yes
B. Sometimes
C. No
D. Don't Know
E. No Response

1983 > 1979 = +
1979 > 1983 = –
No signs are attached to ties
*Significant at the .05 level

| | Percent of Responses | | | | | | | | | | DIRECTION OF DIFFERENCE |
| | 1979 n=754 | | | | | 1983 n=918 | | | | | |
	A	B	C	D	E	A	B	C	D	E	
1. Do you feel that your child is making satisfactory progress in reading?	67	19	10	3	1	70	18	7	4		+
2. Are you satisfied with the reading program in your child's school?	60	11	11	16	2	70	10	5	14		+
3. Would you feel comfortable discussing any reading problem your child might have with his/her teachers?	91	3	3	1	1	92	3	2	2		+
4. Does the school let you know about your child's progress (by report card or teacher conference) in reading as often as you would like?	63	9	25	1	2	76	6	16	2		+
5. Has the school let you know your child's reading grade level?	48	4	43	3	2	53	4	38	4		+
6. Do you feel that your child enjoys learning to read and/or improving his/her reading skills?	67	23	8	2		68	23	8	1		+
7. Does your child think reading is an important skill?	79	12	5	5	1	78	14	4	3		+
8. Do you know what type of reading activities your child does in school?	38	21	37	3	1	43	24	26	5		+
9. Do your child's reading teachers provide additional materials for him/her?	49	17	13	24	2	64	15	11	10		+
10. Does your child read at home?	66	28	5	0	1	64	29	4			0

DIRECTIONS: Choose the appropriate answer for the following statements and mark the correct box.

A. Yes
B. Sometimes
C. No
D. Don't Know
E. No Response

1983 > 1979 = +
1979 > 1983 = -
No signs are attached to ties
*Significant at the .05 level

	Percent of Responses										DIRECTION OF DIFFERENCE
	1979 n=754					1983 n=918					
	A	B	C	D	E	A	B	C	D	E	
11. Do you think reading is an important skill?	99	0	0	0	1	99					0
12. Does your child's reading teacher encourage your child to read at home?	61	10	4	22	2	76	8	3	12		+

*P = .0059

APPENDIX D

STUDENT SURVEY IN READING K-3

PERCENTAGE OF RESPONSES 1979 AND 1982

GUILFORD COUNTY SCHOOL SYSTEM

STUDENT SURVEY IN READING K-3

This survey is part of an attempt to describe the present reading program of the Guilford County Schools. It is to be completed by the student in the presence of an adult. The adult will read the question to the student and record the student's answer.

DIRECTIONS: Choose the appropriate answer for the following statements and mark the correct box.

A. Yes   1983 > 1979 = +
B. Sometimes   1979 > 1983 = -
C. No   No signs are attached to ties
D. Don't Know   NS = Not Significant
E. Not applicable.

	Percent of Responses 1979 n=504					Percent of Responses 1983 n=417					DIRECTION OF DIFFERENCE
	A	B	C	D	E	A	B	C	D	E	
1. Do you like what your teacher plans for you to read every day?	85	11	2	1	0	84	12	3	1	0	0
2. Are you learning to read?	95	2	1	1	0	94	2	3	1	0	−
3. Do you like to read?	86	9	3	1	0	84	11	3	1	0	0
4. Do you read every day in school?	71	16	13	0	0	71	17	12	1	0	+
5. Do you read at home?	61	31	8	0	0	65	28	7	0	0	+

P = .5000 NS

APPENDIX E

STUDENT SURVEY IN READING 4-12

PERCENTAGE OF RESPONSES 1979 AND 1982

GUILFORD COUNTY SCHOOL SYSTEM

STUDENT SURVEY IN READING 4-12

This survey is part of an attempt to describe the present reading program of the Guilford County Schools. It is to be completed by the student in the presence of an adult. The student will read the question and check the appropriate box.

DIRECTIONS: Choose the appropriate answer for the following statements and mark the correct box.

A. Yes
B. Sometimes
C. No
D. Don't Know
E. No Response

1983 > 1979 = +
1979 > 1983 = -
No signs are attached to ties

	Percent of Responses 1979 n=645					Percent of Responses 1983 n=714					DIRECTION OF DIFFERENCE
	A	B	C	D	E	A	B	C	D	E	
1. Do you know which activities you do at school that are supposed to help you read better?	64	20	7	6	3	68	20	7	5		+
2. Do you read at home as part of homework?	45	47	8		1	46	44	11			-
3. Do you read at home for fun?	46	38	16		1	46	37	16			0
4. Does your school prove you with interesting materials to read?	52	38	9	1		61	29	6	4		0
5. Does your school and/or teacher let you know about your progress in reading as often as you like?	37	24	33	6	5	39	28	24	8		+
6. Do you ask your reading/language arts teachers for help if you have problems with reading?	43	32	24	1	1	55	29	15	1		+
7. Does your teacher give you help with reading when you need it?	74	17	6	3	1	80	15	3	2		+
8. Are you satisfied with the help you are receiving?	75	16	5	3	1	81	12	4	2		+
9. Do you feel that your reading has improved this school year?	68	8	12	11	1	69	11	10	9		+
10. Do you enjoy your reading activities at school?	40	49	10		1	91	6	1	1	1	-

GUILFORD COUNTY SCHOOL SYSTEM

STUDENT SURVEY IN READING 4-12

This survey is part of an attempt to describe the present reading program of the Guilford County Schools. It is to be completed by the student in the presence of an adult. The student will read the question and check the appropriate box.

DIRECTIONS: Choose the appropriate answer for the following statements and mark the correct box.

A. Yes   1983 > 1979 = +
B. Sometimes   1979 > 1983 = -
C. No   No signs are attached to ties
D. Don't Know   NS = Not Significant
E. No Response

| | | Percent of Responses | | | | | | | | | DIRECTION OF DIFFERENCE |
| | | 1979 n=645 | | | | | 1983 n=714 | | | | |
		A	B	C	D	E	A	B	C	D	E	
11.	Do you think reading is important?	92	7	1			1	91	6	1	1	−
12.	Are you happy with the way you read?	62	23	12	2	1	64	23	11	2		+

P = .1124 NS

APPENDIX F

ROUND #1 IMPACT QUESTIONNAIRE

GUILFORD COUNTY SCHOOL SYSTEM

ROUND #1 IMPACT QUESTIONNAIRE

DIRECTIONS: Identify the features of your reading program that are practiced more or that
              are practiced less as a result of implementing R-E-A-D in your school. Choose
              the appropriate answer for the following statements and mark the correct box.

              A. This is practiced much more often since implementing the R-E-A-D program.
              B. This is practiced a little more often since implementing the R-E-A-D
                  program.
              C. This practice has not changed since the implementation of R-E-A-D.
              D. This is practiced a little less often since the implementation of
                  R-E-A-D.
              E. This is practiced much less often since the implementation of R-E-A-D.

I. Program Organization

     A. In Organizing Students for Reading Instruction...I

	+2	+1	0	-1	-2
	A	B	C	D	E
1. Divide students into more than one group on the basis of reading ability (i.e., ability grouping).					
2. Group students on the basis of specific reading needs (i.e., special needs grouping).					
3. Base classroom reading instruction on the idea of whole-class grouping (i.e., one group).					
4. Instruct students individually in reading rather than in groups (i.e., individualized reading instruction).					
5. Group students for reading instruction on the basis of commonly-shared interests (i.e., interest grouping).					
6. Use the results of standardized reading achievement tests as measured by the State Annual Testing Program.					
7. Move students from one group to another (i.e., flexible grouping) as needs vary.					
8. Group students on the basis of the group they were in last year.					
9. Organize my reading program with the help of reading specialists.					
10. Use parents, paraprofessionals and/or community resources.					
11. Organize my reading program with the help of colleagues.					

DIRECTIONS: Identify the features of your reading program that are practiced more or that are practiced less as a result of implementing R-E-A-D in your school. Choose the appropriate answer for the following statements and mark the correct box.

A. This is practiced much more often since implementing the R-E-A-D program.
B. This is practiced a little more often since implementing the R-E-A-D program.
C. This practice has not changed since the implementation of R-E-A-D.
D. This is practiced a little less often since the implementation of R-E-A-D.
E. This is practiced much less often since the implementation of R-E-A-D.

	+2	+1	0	-1	-2
	A	B	C	D	E

B. Concerning Goals and Objectives...I

12. Use the school system's curriculum guide in reading as the source of reading program goals.

13. List general goals for the reading program based on the assessment of the students' reading strengths and reading needs.

14. Maintain a record keeping system to keep track of individual progress toward specific objectives.

15. Use the basal reader as the source of reading program goals.

C. In Planning for Skills Instruction...I

16. Organize classroom reading instruction on the basis of skill levels represented in the class.

17. Establish a sequence of reading skills based on assessment of student reading needs.

18. Use the basal reading series to determine the sequence in which reading skills are taught.

19. Select practice activities that match instructional objectives.

20. Organize the classroom with learning centers.

D. When Instructing Students in Reading...I

21. Teach reading through non-basal materials.

22. Assign workbook pages as practice activities which match instructional objectives.

23. Use a file of workbook pages and exercises classified by skill and level.

DIRECTIONS: Identify the features of your reading program that are practiced more or that are practiced less as a result of implementing R-E-A-D in your school. Choose the appropriate answer for the following statements and mark the correct box.

A. This is practiced much more often since implementing the R-E-A-D program.
B. This is practiced a little more often since implementing the R-E-A-D program.
C. This practice has not changed since the implementation of R-E-A-D.
D. This is practiced a little less often since the implementation of R-E-A-D.
E. This is practiced much less often since the implementation of R-E-A-D.

+2	+1	0	-1	-2
A	B	C	D	E

24. Use workbooks as the major guide to introducing reading skills.

25. Have teacher-pupil planning sessions characterized by considerable give-and-take.

26. Prepare a directory of commercial reading materials available within my school.

27. Use the State's Guide for Evaluation of Materials.

28. Code reading materials to reading objectives.

29. Use audio-visual materials.

II. Instructional Practices

A. Methods of Teaching...I

30. Use the basal approach.

31. Use the language experience approach.

32. Use the phonics approach.

33. Use programmed instruction.

34. Use an individualized approach.

35. Use management systems. EXAMPLE: PRI.

36. Use an eclectic approach.

37. Other (list on number 73).

38. Encourage a child to select topics he/she or a group may wish to read about.

39. Release a child from group work to do individual reading.

DIRECTIONS: Identify the features of your reading program that are practiced more or that
are practiced less as a result of implementing R-E-A-D in your school. Choose
the appropriate answer for the following statements and mark the correct box.

    A. This is practiced much more often since implementing the R-E-A-D program.
    B. This is practiced a little more often since implementing the R-E-A-D
       program.
    C. This practice has not changed since the implementation of R-E-A-D.
    D. This is practiced a little less often since the implementation of
       R-E-A-D.
    E. This is practiced much less often since the implementation of R-E-A-D.

III. Teacher Evaluation Practice...!

	+2 A	+1 B	0 C	-1 D	-2 E
40. Use records of independent reading done by each student.					
41. Consider a child's ability to discuss what he/she has heard others read aloud.					
42. Consider work in reading other than the basal program material.					
43. Compare achievement to behavioral objectives by the means of a criterion-referenced test (the objectives state conditions under which a child should do something which teachers can observe, to a degree which the teachers have specified).					
44. Use results of an informal reading inventory.					
45. Consider to what degree a student's textbooks or required reading are matched to his reading level.					
46. The child and his/her teacher make a "performance contract" and the teacher assesses the child's progress in completing this contract.					

IV. Knowledge and Practice

While the purpose of R-E-A-D was not to increase knowledge, it may have been a by-
product. Please indicate areas where knowledge and practices have increased as a
result of the R-E-A-D program.

Use the following choices for questions 47-63.
A. Much, B. Little, C. None

	A	B	C
47. Training in teaching reading.			
group may wish to read about.			
48. Knowledge of the diagnostic-prescriptive approach to reading.			

DIRECTIONS: Identify the features of your reading program that are practiced more or that are practiced less as a result of implementing R-E-A-D in your school. Choose the appropriate answer for the following statements and mark the correct box.

A. Much
B. Little
C. None

	A	B	C
49. Use of a diagnostic prescriptive approach to teaching reading.			
50. Knowledge of motivation techniques.			
51. Use of these techniques.			
52. Knowledge of our school system's curriculum guide.			
53. Use of this guide.			
54. Knowledge of management system techniques.			
55. Use of a management system.			
56. Knowledge of standardized norm-referenced tests.			
57. Use of the results of standardized norm-referenced tests.			
58. Knowledge of criterion-referenced tests.			
59. Use of the results of criterion-referenced tests.			
60. Knowledge of criteria for selecting materials.			
61. Use of criteria for selecting materials.			
62. Knowledge of the sources for reading materials.			
63. Knowledge of the issue of accountability and its implications.			

V. Teacher's Evaluation of the Reading Program

64. How successful is your school's reading program on the whole?
    A. Much more, B. A little more, C. No change

65. How satisfied are you with the reading skills and habits of the students in your classes?
    A. Much more, B. A little more, C. No change

	A	B	C
66.			
67.			

66. How effective do you find the present reporting system for record keeping of reading skills as students move from grade to grade?
A. Much more effective, B. A little more effective,
C. No change

67. How effective do you find the present reporting system in reading for school to school transfer?
A. Much more effective, B. A little more effective,
C. No change

ABOVE AVERAGE	AVERAGE	BELOW AVERAGE	ALL
A	B	C	D

68. The reading program in your school is geared more toward which group of students since implementing the R-E-A-D program?

69. What evidence do you consider in making your judgment as to the instructional levels of the children since implementing the R-E-A-D program?

_____ Informal Reading Inventory   _____ Individual Conference
_____ Publisher's tests   _____ Own tests   _____ Books pupil chooses to read
_____ Oral Reading   _____ Workbooks   _____ Standardized tests

APPENDIX G

ROUND #2 IMPACT QUESTIONNAIRES

DIRECTIONS

QUESTIONNAIRES:

READING CURRICULUM COMMITTEE

SCHOOLS:

411

320

368

334

360

336

READING CURRICULUM COMMITTEE

GUILFORD COUNTY SCHOOL SYSTEM

Round #2 Impact Questionnaire

DIRECTIONS

A. Distribute these questionnaires to the same teachers who
    completed the Round #1 Impact Questionnaires.

B. Bring all questionnaires to the meeting on April 6.

Thank you.

GUILFORD COUNTY SCHOOL SYSTEM

IMPACT QUESTIONNAIRE NUMBER TWO

READING CURRICULUM COMMITTEE

The following possible outcome impact statements have been identified by an analysis of questionnaire number one. Please review each statement and rank order the ten that relate most closely to the R-E-A-D curriculum. 1 = most closely related, 2 = second most closely related, etc. If you want to add your comments, please use the space provided. Please return on April 6, 1983.

Rank 1-10	Impact Statement	Comments
2	Teacher's knowledge of the diagnostic-prescriptive approach to reading has increased.	
1	Teacher' knowledge of the school system's curricular has increased.	
6	Teacher's use of the school system's curricular guide has increased.	
5	The reading program is more successful.	
4	Teachers are more satisfied with the reading progam.	
10	Principals are more satisfied with the reading program.	
3	Student achievement in reading has increased.	
9	Students are more satisfied with the reading program.	

Rank 1-10	Impact Statement	Comments
8	The record keeping of reading skills is more effective.	
7	Use of the system's curricular guide as the source of reading program goals has increased.	

GUILFORD COUNTY SCHOOL SYSTEM

IMPACT QUESTIONNAIRE NUMBER TWO

SCHOOL 411

The following possible outcome impact statements have
seen identified by an analysis of questionnaire number one.
Please review each statement and rank order the ten that
relate most closely to the R-E-A-D curriculum. 1 = most
closely related, 2 = second most closely related, etc. If
you want to add your comments, please use the space
provided.

Rank 1-10	Impact Statement	Comments
_____	Student achievement in reading has improved.	_____ _____
_____	Students are more satisfied with the reading program.	_____ _____
_____	Parents are more satisfied with the reading program.	_____ _____
_____	Teachers are more satisfied with the reading program.	_____ _____
_____	Principals are more satisfied with the reading program.	_____ _____
_____	Teacher's knowledge of the school system's curricular guide has increased.	_____ _____ _____
_____	Teacher's use of the school system's curricular guide has increased.	_____ _____ _____
_____	The record keeping of reading skills as students move from grade to grade is a little more effective.	_____ _____ _____ _____
_____	The record keeping of reading skills as students transfer from school to school is a little more effective.	_____ _____ _____ _____

Rank 1-10	Impact Statement	Comments
_____	Grouping students on the basis of specific reading needs has increased.	_____
_____	General goals for the reading program based on the assessment of the students' strengths and needs has increased.	_____
_____	More classroom reading instruction is based on the skills levels represented in the class.	_____
_____	A sequence of reading skills based on assessment of student's reading needs is established.	_____
_____	Reading is taught through non-basal materials.	_____
_____	Training in reading has increased.	_____
_____	Use of a diagnostic-prescriptive approach to reading has increased.	_____
_____	Knowledge of motivation techniques has increased.	_____
_____	Knowledge of motivation techniques has increased.	_____
_____	Knowledge of standardized norm-referenced tests has increased.	_____
_____	Knowledge of the sources of reading materials has increased.	_____
_____	Knowledge of the issue of accountability and its implications has increased.	_____

GUILFORD COUNTY SCHOOL SYSTEM

IMPACT QUESTIONNAIRE NUMBER TWO

SCHOOL 320

The following possible outcome impact statements have been identified by an analysis of questionnaire number one. Please review each statement and rank order the ten that relate most closely to the R-E-A-D curriculum. 1 = most closely related, 2 = second most closely related, etc. If you want to add your comments, please use the space provided.

Rank 1-10	Impact Statement	Comments
_____	Student achievement in reading has improved.	_____ _____
_____	Students are more satisfied with the reading program.	_____ _____
_____	Parents are more satisfied with the reading program.	_____ _____
_____	Teachers are more satisfied with the reading program.	_____ _____
_____	Principals are more satisfied with the reading program.	_____ _____
_____	Teacher's knowledge of the school system's curricular guide has increased.	_____ _____ _____
_____	Teacher's use of the school system's curricular guide has increased.	_____ _____ _____
_____	The record keeping of reading skills as students move from grade to grade is a little more effective.	_____ _____ _____
_____	The record keeping of reading skills as students transfer from school to school is a little more effective.	_____ _____ _____

Rank 1-10	Impact Statement	Comments
_____	Use of the school system's curricular guide in reading as the source of reading program goals.	_____ _____ _____ _____
_____	List general goals for the reading program based on the assessment of the students' reading strengths and reading need.	_____ _____ _____ _____
_____	Use of a diagnostic-prescriptive approach to teaching reading has increased.	_____ _____ _____ _____
_____	Knowledge of standardized norm-referenced tests has increased.	_____ _____ _____
_____	Use of the results of standardized norm-referenced tests has increased.	_____ _____ _____
_____	Knowledge of the issue of accountability and its implications has increased.	_____ _____ _____
_____	The school's reading program is more successful.	_____ _____
_____	Teachers are more satisfied with the reading skills and habits of the students.	_____ _____ _____
_____	Oral reading as an indicator of the student's instructional level has increased.	_____ _____ _____
_____	Use of teacher's own test as an indicator of the student's instructional level has increased.	_____ _____ _____

GUILFORD COUNTY SCHOOL SYSTEM

IMPACT QUESTIONNAIRE NUMBER TWO

SCHOOL 368

The following possible outcome impact statements have seen identified by an analysis of questionnaire number one. Please review each statement and rank order the ten that relate most closely to the R-E-A-D curriculum. 1 = most closely related, 2 = second most closely related, etc. If you want to add your comments, please use the space provided.

Rank 1-10	Impact Statement	Comments
_____	Student achievement in reading has improved.	_____
_____	Students are more satisfied with the reading program.	_____
_____	Parents are more satisfied with the reading program.	_____
_____	Teachers are more satisfied with the reading program.	_____
_____	Principals are more satisfied with the reading program.	_____
_____	Teacher's knowledge of the school system's curricular guide has increased.	_____
_____	Teacher's use of the school system's curricular guide has increased.	_____
_____	The record keeping of reading skills as students move from grade to grade is a little more effective.	_____
_____	The record keeping of reading skills as students transfer from school to school is a little more effective.	_____

Rank 1-10	Impact Statement	Comments

_____	Teachers organize the reading program with the help of colleagues.	_____ _____
_____	Use of the Informal Reading Inventory as an indicator of the student's instructional level has increased.	_____ _____ _____ _____
_____	Knowledge of management system techniques has increased.	_____ _____ _____
_____	Use of management system techniques has increased.	_____ _____
_____	Use of standardized norm-referenced test results has increased.	_____ _____ _____
_____	Use of the results of criterion-referenced tests has increased.	_____ _____ _____
_____	Use of criteria for selecting materials has increased.	_____ _____
_____	Knowledge of the sources of reading materials has increased.	_____ _____ _____
_____	Knowledge of the issue of accountability and its implications has increased.	_____ _____ _____
_____	The reading program is geared more toward meeting the needs of all of the student.	_____ _____ _____

GUILFORD COUNTY SCHOOL SYSTEM

IMPACT QUESTIONNAIRE NUMBER TWO

SCHOOL 334

The following possible outcome impact statements have been identified by an analysis of questionnaire number one. Please review each statement and rank order the ten that relate most closely to the R-E-A-D curriculum. 1 = most closely related, 2 = second most closely related, etc. If you want to add your comments, please use the space provided.

Rank 1-10	Impact Statement	Comments
_____	Student achievement in reading has improved.	_____
_____	Students are more satisfied with the reading program.	_____
_____	Parents are more satisfied with the reading program.	_____
_____	Teachers are more satisfied with the reading program.	_____
_____	Principals are more satisfied with the reading program.	_____
_____	Teacher's knowledge of the school system's curricular guide has increased.	_____
_____	Teacher's use of the school system's curricular guide has increased.	_____
_____	The record keeping of reading skills as students move from grade to grade is a little more effective.	_____
_____	The record keeping of reading skills as students transfer from school to school is a little more effective.	_____

Rank 1-10	Impact Statement	Comments

<u>Rank 1-10</u>   <u>Impact Statement</u>   <u>Comments</u>

_____ Use of the school system's
curricular guide in reading
as a source of reading
program goals.

_____ Use of a diagnostic-
prescriptive approach to
reading has increased.

_____ Use of a management system
has increased.

_____ Knowledge of the issues of
accountability and its
implications has increased.

_____ Use of an Informal Reading
Inventory as an indicator of
the student's instructional
level has increased.

_____ Use of Standardized tests as
an indicator of the student's
instructional level has
increased.

_____ Use of the teacher's own test
as an indicator of the
student's instructional level
has increased.

GUILFORD COUNTY SCHOOL SYSTEM

IMPACT QUESTIONNAIRE NUMBER TWO

SCHOOL 360

The following possible outcome impact statements have seen identified by an analysis of questionnaire number one. Please review each statement and rank order the ten that relate most closely to the R-E-A-D curriculum. 1 = most closely related, 2 = second most closely related, etc. If you want to add your comments, please use the space provided.

Rank 1-10	Impact Statement	Comments
_____	Student achievement in reading has improved.	
_____	Students are more satisfied with the reading program.	
_____	Parents are more satisfied with the reading program.	
_____	Teachers are more satisfied with the reading program.	
_____	Principals are more satisfied with the reading program.	
_____	Teacher's knowledge of the school system's curricular guide has increased.	
_____	Teacher's use of the school system's curricular guide has increased.	
_____	The record keeping of reading skills as students move from grade to grade is a little more effective.	
_____	The record keeping of reading skills as students transfer from school to school is a little more effective.	

Rank 1-10	Impact Statement	Comments
_____	Use of the school system's curricular guide in reading as the source of reading program goals.	_____
_____	Knowledge of management system techniques has increased.	_____
_____	Knowledge of criteria for selecting materials has increased.	_____
_____	Knowledge of the sources of reading materials has increased.	_____

GUILFORD COUNTY SCHOOL SYSTEM

IMPACT QUESTIONNAIRE NUMBER TWO

SCHOOL 336

The following possible outcome impact statements have been identified by an analysis of questionnaire number one. Please review each statement and rank order the ten that relate most closely to the R-E-A-D curriculum. 1 = most closely related, 2 = second most closely related, etc. If you want to add your comments, please use the space provided.

Rank 1-10	Impact Statement	Comments
_____	The reading program is a little more successful.	_____ _____
_____	Teachers are a little more satisfied with the reading program.	_____ _____
_____	The present reporting system for record keeping of reading skills as students move from grade to grade is a little	_____ _____ _____
_____ _____	The present reporting system in reading for school to school transfer is a little more effective.	_____ _____ _____
_____	The Informal Reading Inventory as an indicator of the student's instructional level has increased.	_____ _____ _____
_____	Student achievement in reading has improved.	_____ _____
_____	Parents are more satisfied with the reading program.	_____ _____
_____	Principals are more satisfied with the reading program.	_____ _____
_____	Teachers are more satisfied with the reading program.	_____ _____

Rank 1-10	Impact Statement	Comments
_____	Teacher's knowledge of the diagnostic-prescriptive approach to teaching reading has increased.	_____ _____ _____
_____	Teacher's knowledge of the school system's curricular guide has increased.	_____ _____ _____
_____	Teacher's use of the school system's curricular guide has increased.	_____ _____
_____	Students are more satisfied with the reading program.	_____ _____