DOCUMENT RESUME

ED 027 165 24 RE 001 560

By-Niles, Olive S.
Evaluation of Three Methods of Teaching First Grade Reading to Children Likely to Have Difficulty with Reading.
Massachusetts State Dept. of Education, Boston.
Spons Agency-Office of Education (DHEW), Washington, D.C. Bureau of Research.
Report No-CRP-2702
Bureau No-BR-5-0469
Pub Date 65
Contract-OEC-5-10-084
Note-107p.
EDRS Price MF-$0.50 HC-$5.45
Descriptors-Attitudes, Basic Reading, *Beginning Reading, *Childrens Books, Grade 1, Reading Instruction, *Reading Readiness, *Reading Research, *Remedial Reading Programs, Remedial Teachers

Forty first-grade classes were divided into four treatment groups to determine the effectiveness of three reading methods with low ability students. Treatment A subjects used the regular basal program. Treatment B subjects used the same basal program, with the low subgroup receiving additional instruction from remedial reading teachers. Treatment C low subgroups used Houghton Mifflin readiness materials and tradebooks, while the rest of the class used the basal program. Treatment D low subgroups received additional remedial instruction as well as regular classroom instruction with the Houghton Mifflin readiness materials and tradebooks. At the end of the 140-day experimental period, the Stanford Achievement Test, Primary 1, Form X, and the San Diego Attitude Inventory were administered to all subjects. Additional testing was done with a random sample. Fifty-three tables present and compare the data, and eight tables summarize the trends. It was tentatively concluded that use of the Houghton Mifflin materials plus tradebooks instead of basal readers was more effective than either a regular basal program or remedial teacher time spent with the low subgroups. However, the combination of the special materials with remedial teacher time appeared to be more effective than either by itself. (CM)

EVALUATION OF THREE METHODS OF TEACHING FIRST GRADE READING TO CHILDREN LIKELY TO HAVE DIFFICULTY WITH READING

COOPERATIVE RESEARCH PROJECT 2702

Massachusetts State Department of
Education, Newbury Street, Boston,
Massachusetts, in cooperation with
the Springfield Public Schools,
Springfield, Massachusetts

Project Director: Olive S. Niles

1965

The research reported herein was supported by the Cooperative Research Program of the Office of Education, U. S. Department of Health, Education, and Welfare.

TABLE OF CONTENTS

ACKNOWLEDGEMENTS

A project of this type requires the cooperation of many persons.

Most important for its success were the forty teachers of the first grade classrooms involved in the study. With very little warning or preparation, they devoted themselves to the task with a commendable professional spirit.

The administration of the Springfield School System also cooperated wholeheartedly: the superintendent and assistant superintendent in charge of elementary education by lending approval and support, and the principals of the elementary schools by helping to see that arrangements were carried out as planned.

The research department of the Springfield Public Schools was helpfully involved in making the proposal and studying the results.

Members of the Massachusetts State Department of Education in Boston cooperated in the process of formulating the original plan, following through with the financial arrangements, and helping with statistical procedures.

Mrs. Mildred Lowe, who served as coordinator of the whole study, and Miss Agnes McCarthy and Mrs. Helen Thayer, who were the traveling remedial teachers, gave far more than their assigned time and energy to making the program work.

LIST OF TABLES

The Problem

Children with good academic ability and with physical and emotional maturity learn to read by almost any method, though admittedly more efficiently with some methods than with others.

It is the children with less than average mental ability and/or poor physical and emotional maturity and, in many cases, disadvantaged home situations who find the first-grade reading program so frustrating. They meet failure at so young an age that "reading" may be forevermore a bad word to them.

This study was concerned mainly with trying to find out:

1) What predictive tests are useful and practical in identifying at the very beginning of their first-grade experience those children likely to have trouble in learning to read

2) Which of the four procedures tested in the study is most effective in teaching the children so identified (to be designated as the "low subgroups") to read

A further question raised in the study was as follows:

3) What, if any, differences occurred in the performance of the <u>total</u> group of children in the classrooms in which the four kinds of procedures were being carried on with children in the "low subgroups"?

Null Hypotheses

1. There is no significant difference in the distribution of test scores at the end of grade one among the "low subgroups" within the total treatment groups in each of the four situations (one control and three experimental).
2. There is no significant difference in the distribution of test scores at the end of grade one among the total treatment groups in each of the four situations.

Related Research

Predicting success in first-grade reading

Many factors are associated with success in first-grade reading. Chronological age, however, has been shown to have either a very low or even a negative correlation with reading achievement at this level. This factor, therefore, cannot be used as a predictor.

Correlations of scores on intelligence tests with measures of reading achievement tend to fall between .35 and .65, with the lower correlations generally obtained from studies of young children in grades 1 or 2. (1) Manolakes and Sheldon, in a study of children in grades one to twelve, found a higher correlation between intelligence and reading achievement after fourth grade. (2)

Durkin's report of 49 children who learned to read before they entered grade 1 showed an M. A. range from 5.1 to 10.7. She states: "Current

---

(1) Guy L. Bond and Miles A. Tinker, Reading Difficulties: Their Diagnosis and Correction. New York: Appleton-Century-Crofts, 1957, p. 42

(2) George Manolakes and William D. Sheldon, "The Relation Between Reading-Test Scores and Language-Factors Intelligence Quotients," Elementary School Journal, Vol. 55 (Feb. 1955), pp. 346-350

intelligence tests are seriously inadequate in identifying and measuring 'what it takes' to learn to read. The fact that over one-third of the children in this study had IQ's less than 110 would at least suggest this test inadequacy, and would also suggest the existence of important intellectual factors or abilities not included in the IQ's." (3)

Experiments by Gates showed that the relationship between reading achievement and mental age varied with the kind of instructional procedures used. (4) This conclusion was confirmed by Roslow (5) and by several subsequent studies. The value of an intelligence test as a general predictor of success in first-grade reading is therefore in serious doubt.

That the low correlations found between mental age and early reading success may be partly the result of the type of intelligence test used in the investigations is a possibility. A few studies of early readers have seemed to challenge the conclusion that intelligence does, in fact, have a low relationship to early ability to read. For example, the Plessas and Oakes study (6) reported a mean Wisc score of 128 for twenty children who could read above 2.0 on the California Reading Test in December of first grade. However, this same study concluded that besides having su-

---

(3) Dolores Durkin, "Children Who Read Before Grade One," The Reading Teacher, Vol. 14 (Jan. 1961), pp. 163-166

(4) Arthur I. Gates, "The Necessary Mental Age for Beginning Reading," Elementary School Journal, Vol. 37 (March 1937), pp. 397-408

(5) Sydney Roslow, "Reading Readiness and Reading Achievement in the First Grade," Journal of Experimental Education, Vol. 9 (Dec. 1940), pp. 154-159

(6) Gus P. Plessas and Clifton R. Oakes, "Prereading Experiences of Selected Early Readers," The Reading Teacher, Vol. 17 (Jan. 1964), pp. 241-245

perior intelligence these children were surrounded by a very positive environment geared to interest in language and that they had been taught to read by someone in the family. Their early progress was not a "chance happening." Their environment was of the type which stimulates perceptual growth. This finding supports observations of the Durkin study of early readers. (7)

Durkin reported that the children who learned to read before they entered first grade had certain personal characteristics in common: good memories, ability to concentrate, curiosity, persistence, and self-reliance. These personality traits upon which Durkin places emphasis are hard, if not impossible, to measure with a standardized instrument. Reading readiness tests, as they exist today, are affected by these traits but do not measure them directly. Spache states the need for more knowledge in this area: "Long-range studies of the relationship between personality traits of primary children and ultimate reading success are vitally needed." (8) The lack of such instruments may, indeed, make accurate prediction impossible at this time.

Many attempts, however, have been made to measure specific skills which are associated with readiness to read and to find a combination of such measures with predictive value. General agreement exists that a child must be skilled in visual perception of letters and combinations of letters. Vernon points out the kinds of visual perception difficulties children may have and discusses the effect of these difficulties on

---

(7) Op. cit.

(8) George D. Spache, Toward Better Reading, Garrard Press, 1963, p. 12

learning to read. (9) Goins found evidence of the importance of the ability to see and keep in mind both a perceptual whole and the parts within it. (10) There is some evidence that many of the older-type tests and practice exercises in gross visual discrimination involving pictures and geometrical forms are, for all but a few children, a waste of time. Durrell makes this statement: "All children (of about 2000 entering first grade) were able to match capital letters as well as lower-case letters. Exercises in this ability should be omitted from reading readiness materials. It appears to follow that matching of non-word forms and pictures as preliminary instruction for letter and word perception is relatively useless." (11) Contrary evidence is given by Goins (12) who found that tests involving visual discrimination of geometric figures seemed to have value as predictors of first-grade reading achievement.

Barrett conducted an extensive study of seven visual discrimination tests as predictors. Of these, the best single predictor was the Gates Reading Letters and Numbers Test. Next in order were a Pattern Copying Test and the Gates Word Matching Test. Barrett cautions against assuming cause and effect relationships. "In other words such an ability(as reading letters, etc.) may be a symptom of many kinds of experiences with letters, numbers, words, and stories; therefore, it

---

(9) M. D. Vernon, Backwardness in Reading, Cambridge University Press, 1957

(10) Jean Turner Goins, Visual Perceptual Abilities and Early Reading Progress, Supplementary Education Monographs, No. 87, University of Chicago Press, February 1958

(11) Donald D. Durrell, ed., "Success in First Grade Reading," Journal of Education, Boston University, Vol. 140 (Feb. 1958), p. 5

(12) Op. cit.

should not be assumed from this study that success in first-grade reading will be insured by simply teaching children to discriminate, recognize and name letters and numbers." Barrett also states that the use of the visual discrimination tasks imposed in his study "did not provide enough prediction precision to warrant their use alone in predicting first-grade reading achievement for individuals." (13)

Auditory factors have consistently been shown to be of great importance in predicting success in first grade reading. Sister Mary Nila found the four chief factors in early reading success to be auditory discrimination, visual discrimination, range of information, and mental age in that order. (14) Harris stresses the lack of auditory perception among youngsters who have difficulty with reading and notes that children who seem to have normal hearing on the usual hearing tests often cannot distinguish small differences in sounds or have difficulty even in hearing the sounds of separate letters within words. (15)

Thompson reports a study of 105 children completing grades 1 and 2. These children were given three tests of auditory discrimination: the Wepman Test of Auditory Discrimination, the Boston University Speech Sound Discrimination Test, and the Auditory Discrimination and Orienta-

---

(13) Thomas C. Barrett, "Visual Discrimination Tasks as Predictors of First Grade Reading Achievement," The Reading Teacher, Vol. 18 (Jan. 1965), pp. 276-282

(14) Sister Mary Nila, O.S.F., "Foundations of a Successful Reading Program," Education, Vol. 73 (May 1953), pp. 543-555

(15) Albert J. Harris, How to Increase Reading Ability, Longmans, Green, 1961, p. 230

tion subtests of the SRA Reading Analysis. Thompson found that auditory discrimination ability in first grade is "highly prognostic" of later reading success. (16)

In general, existing readiness test batteries have not proved to be good predictors. Karlin, for example, studied the Metropolitan Readiness Test as a predictor of scores on the Gates Primary Reading Test, Type Three, Paragraph Reading. When chronological age and intelligence were held constant, the correlation was .25. (17) Berwick found the correlation between reading achievement and the Lee-Clark Reading Readiness Test to be .47. (18) Mattick, working with 972 children, found scores on the Metropolitan Readiness Test to have a higher correlation with achievement in first grade reading (as rated by teachers at the end of October) than kindergarten teachers' judgment. However, the kindergarten teachers' judgment was superior to scores on the California Short-Form Test of Mental Maturity, the Lee-Clark Reading Readiness Test, and the Lorge-Thorndike Intelligence Tests. (19)

---

(16) Bertha B. Thompson, "A Longitudinal Study of Auditory Discrimination," Journal of Educational Research, Vol. 56 (March 1963), pp. 376-8

(17) Robert Karlin, "The Prediction of Reading Success and Reading Readiness Tests," Elementary English, Vol. 34 (May 1957), pp. 320-322

(18) Mildred M. Berwick, An Evaluation of the Prognostic Value of Certain Pre-reading Tests for Reading Achievement, Unpublished Master's Thesis, Boston University, 1947

(19) William E. Mattick, "Predicting Success in the First Grade," Elementary School Journal, Vol. 63 (Feb. 1963), pp. 273-276

Powell and Parsley found a low relationship between individual scores on the California Reading Test, administered in grade 2, and the Lee-Clark Reading Readiness Test given in grade 1. (20)

Gunderson states that most reading readiness tests "are used more effectively as instruments for determining the educational needs of the individual child, so that proper teaching may be planned, than as predictors of achievement." (21)

This failure to predict may be partly because so many of these reading readiness tests have not emphasized three factors which recent research has shown to be of great importance in prediction: the child's ability to identify sounds in spoken words, the level of his letter knowledge, and his learning rate. September level of letter knowledge as a readiness factor was investigated by Gavel in a study of 1506 children. All of her tests of letter names correlated higher with June reading achievement tests than did mental age, ranging from an r of .60 for a test of writing letters dictated to .54 for naming lower case letters. (22)

Durrell suggests that the best way to find out whether a child is ready for reading and will be successful in the process is to

---

(20) Marvin Powell and Kenneth Parsley, "The Relationships Between First Grade Reading Readiness and Second Grade Reading Achievement," Journal of Educational Research, Vol. 54 (Feb. 1961), pp. 229-233

(21) Doris V. Gunderson, Research in Reading Readiness, U. S. Department of Health, Education and Welfare, Office of Education, OE-30013, Bulletin 1964, No. 8, p.24

(22) Sylvia R. Gavel, Patterns of Growth in First Grade Reading, Unpublished doctoral dissertation, Boston University, 1957

teach him some words and see whether he remembers them. (23) The value of a learning rate test as a predictor is shown in the Gavel study (24) in which the correlation of September learning rate tests with June achievement is .51, just below the correlation between June tests of reading achievement and September tests of letter knowledge.

Auditory training during the readiness period

The effectiveness of an emphasis on auditory training has been shown in several studies. An early study of this type was done by Murphy. (25) The Denver study of about 4000 children used materials stressing letter names and sounds and contextual clues. Children who had this program in kindergarten with a follow-up of the same kind of training in grade 1 achieved better in grade 1 than control groups and also better than children who began the same kind of program in grade 1. Differences were significant beyond the .001 level of confidence. (26) Gavel found that February tests with correlations above .60 with June achievement were tests of hearing sounds in words, applied phonics, and ability to give the sound of lower-case letters. (27)

Olson measured the effects of early teaching of letter sounds and

---

(23) Donald D. Durrell, Improving Reading Instruction, World Book, 1956, pp. 49-51

(24) Op. cit.

(25) Helen A. Murphy, An Evaluation of the Effect of Specific Training in Auditory and Visual Discrimination on Beginning Reading, Unpublished doctoral dissertation, Boston University, 1943

(26) Joseph Brzeinski, "Beginning Reading in Denver," The Reading Teacher, Vol. 18 (October 1964), pp. 16-21

(27) Sylvia R. Gavel, "June Reading Achievements of First-Grade Children," Journal of Education, Boston University, Vol. 140 (Feb. 1958), pp. 37-43

names. He concluded that early teaching of various aspects of phonics is essential to rapid progress in reading. He contends that there is no support for the idea that a sight vocabulary should be established before word analysis instruction is begun. (28)

Use of Trade Books in place of basal readers

The contents of basal reading books have been under attack for some time. They are often called "inane," and "superficial," with content subservient to vocabulary control. In the Harvard Report of 1963, Austin has this to say of the matter: "Regardless of the degree of use of the basals, many of the teachers and administrators interviewed during the field study were highly critical of the content of most series. They felt that basals should 'provide a richer literary fare for youngsters. As it is we take little stories and beat them dry when there is nothing to begin with.' Others thought that the content of basal readers, while demanding enough in terms of levels of difficulty, was not sufficiently challenging and bore little relation to the realities of children's lives, particularly in the case of boys. Those who defend the content of basal readers do so on the grounds that, while the stories may be boring to adults, the children like them and that broadening of interests can take place through independent reading." (29)

If the criticism of the content of the basals is true from the child's point of view, it may be particularly important, since what a child reads with interest is likely to lead to better and to more reading. That the

---

(28) Arthur V. Olson, "Growth in Word Perception Abilities as It Relates to Success in Beginning Reading," Journal of Education, Vol. 140 (Feb. 1958), pp. 25-36

(29) Mary C. Austin and Coleman Morrison, The First R, the Harvard Report on Reading in Elementary Schools, Macmillan, 1963, p. 55

basals may not have tapped the breadth of interests first graders actually have is suggested by Byres in her study of the interests of first graders as they are revealed in "show and tell" sessions. (30)

Gans states the point as follows: "... to those who have kept close to the pulse of young children's living, who have noticed their enjoyment of good books and imaginative toys, who have listened to their dramatic play and to their use of the latest language from television and adult happenings, the vocabulary-controlled language and ideas of primary materials are long overdue for a change." (31)

The research literature surveyed revealed no account of the use of trade books to replace basals in group teaching, though there are numerous studies of individualized reading with the use of trade books. In these studies of individualized reading, the evidence is not clearcut that either the individualized program or the basal program is the more successful. (32)

Other factors in the study

No experimental evidence was found on the use of a second, or remedial team-teacher to work with the potential problem readers within the framework of the regular first grade classroom. Also, the combination of the Houghton Mifflin readiness program, with its strong auditory emphasis, with the use of trade books to enhance motivation and effectiveness of teaching with potentially problem readers seems not to have been tested.

---

(30) Loretta Byres, "Pupils' Interests and the Content of Primary Reading Texts," The Reading Teacher, Vol. 17 (January 1964), pp. 227-233

(31) Roma Gans, Common Sense in Teaching Reading, Bobbs-Merrill, 1963, p. 126

(32) George B. Spache, Toward Better Reading, Garrard, 1963, pp. 150-165

PROCEDURES

Locale of the study

The study was conducted in the public schools of Springfield, Massachusetts. This is a city of 174,463 population (1960 census). Since the elementary schools are neighborhood schools, certain schools are located in economically more favored areas than others.

The range of median income is from $2001-$3000 in the neighborhood of one of the schools in the study to $7001-$8000 in the neighborhood of other schools. The range in median number of school years completed by adults is from 9 years in certain neighborhoods to 13 years in others.

Springfield is mostly an industrial city. In addition, there is one large insurance company and several smaller ones. Several colleges are located in or near the city. The public library system is an excellent one with a large central building and seven neighborhood branches. There is also a large museum complex containing historical, science and art museums. The city has forum and concert series which have a long history of excellence. The adult education program serves about 5000 people during any one enrollment period and offers courses in about 125 different subjects.

At the time of the study, no elementary school had a central school library. A small supply of books formed classroom libraries.

Pupil population of the study

There are 38 public elementary schools in Springfield. Thirty-two of these were involved in the first grade study. Of the schools not involved, two were eliminated because they are using a different basal program from the other 36 schools; two, because the enrollment is very small; and two because the school population was, at the time of the study, in a

state of flux.

Among the 32 schools involved in the study there are 111 first grade classes, including combination first and second grades. The largest number in any one school is five; the smallest, one. Mean class size and range of class size for each treatment group in the study is shown in Table I. Class size is based on enrollment as of October 1, 1965. The rather wide range is

TABLE I

Mean and Range of Class Size for Four Treatment Groups

Group	Mean	Range
A	30.2	25-34
B	28.6	25-35
C	28.7	24-31
D	30.0	26-34

accounted for mainly by a deliberate policy of keeping pupil-teacher ratio as low as possible in certain schools where socially handicapped children predominate.

Children are permitted to enter first grade if they have attained an age of 5 years and 7 months by September 10 of the year of entrance. Kindergartens are a part of the regular educational program in Springfield. A check of the amount of pre-first-grade experience of the children in each of the treatment groups indicated that practically all children in each of the groups had participated in a full year of half-time kindergarten (morning or afternoon).

Membership in the first grades of the city is heterogeneous. When

more than one first grade class is present in a school, the principal of the building uses the child's kindergarten experience and the evaluation of the children by kindergarten teachers to group the children for first grade in such a way that all levels of ability are usually found within each first grade class.

Table II shows the number of children in each treatment group with the amount of attrition during the school year. Children repeating grade one were not used in the study.

TABLE II

Number of Children Involved in Each Treatment Group with Amount of Attrition in Each Group

Group	Initial Enrollment	Attrition		
		Repeating Grade 1	Moved	Incomplete Test Data
A	300	33	7	23
B	286	28	25	13
C	286	28	19	19
D	297	36	27	12

Table III shows the mean and standard deviation of the chronological age (in months, as of October 1, 1964) of children in each of the four treatment groups.

TABLE III

Mean and Standard Deviation of Chronological Age in the Four Treatment Groups

Group	N	Mean	S.D.
A	237	73.81	4.1234
B	220	73.84	3.7394
C	219	74.27	3.9301
D	223	74.13	4.1429

Sex of children in the study is summarized in Table IV.

TABLE IV

Sex of Children in the Four Total Treatment Groups

Group	Boys	Girls
A	120	117
B	106	114
C	122	98
D	109	113

The unexpected preponderance of boys in Group C may be significant. It was not intended in the planning of the groups.

Because the elementary schools of Springfield are neighborhood schools with great variations (as shown above) in the background factors of the different neighborhoods, it was necessary to find an objective

means of subdividing the total number of 40 classrooms into the four treatment groups.

Since citywide testing of pupils in first grade has not been a practice in Springfield, it was necessary to equate schools on the basis of results from citywide testing in grades three and five, the assumption being that differences among first grade groups could be estimated from differences in grades 3 and 5. Therefore schools were placed in rank order distributions on the basis of the California Test of Mental Maturity and the reading comprehension subtest of the Iowa Test of Basic Skills in grades 3 and 5 for the year 1963-64. An average of the four rankings (third and fifth grade reading comprehension; third and fifth grade IQ) was used to subdivide the schools into eight groups. This process is summarized in Table V.

TABLE V

Number of Classrooms in Each Treatment Group Selected from Each of the Eight Rank Order Groupings

Rank Order Groupings*	Treatment Group A	Treatment Group B	Treatment Group C	Treatment Group D
1	1	1	1	1
2	2	2	2	2
3	2	2	2	2
4	1	1	1	1
5	1	1	1	1
6	1	1	1	1
7	1	1	1	1
8	1	1	1	1

*Groups of schools are listed in descending order of ability and achievement.

The degree to which these preliminary procedures were successful in equalizing the four treatment groups was checked by examining selected preliminary test scores of children in each treatment group. Tables VI through IX summarize these data for the total treatment groups:

TABLE VI

Comparison of Mean Scores of Total Treatment Groups on the Pintner-Cunningham Primary Test

Group	N	Mean	S.D.	Diff. of Means	C.R.
A	237	35.30	8.53	.69	.861
B	220	35.99	8.41		
A	237	35.30	8.53	1.58	1.868
C	219	36.88	9.38		
A	237	35.30	8.53	1.79	2.096
D	223	33.51	9.73		
B	220	35.99	8.41	.89	1.046
C	219	36.88	9.38		
B	220	35.99	8.41	2.48	2.865
D	223	33.51	9.73		
C	219	36.88	9.38	3.37	3.701
D	223	33.51	9.73		

TABLE VII

Comparison of Mean Scores of Total Treatment Groups on the Murphy-Durrell Phonemes Subtest

Group	N	Mean	S.D.	Diff. of Means	C.R.
A	237	23.18	13.35	.17	.136
B	220	23.35	13.85		
A	237	23.18	13.35	3.90	3.063
C	219	27.08	13.80		
A	237	23.18	13.35	2.08	1.694
D	223	21.10	13.02		
B	220	23.35	13.85	3.73	2.824
C	219	27.08	13.80		
B	220	23.35	13.85	2.25	1.766
D	223	21.10	13.02		
C	219	27.08	13.80	5.98	4.687
D	223	21.10	13.02		

TABLE V111

Comparison of Mean Scores of Total Treatment Groups on the Murphy-Durrell Letter Names Subtest

Group	N	Mean	S.D.	Diff. of Means	Critical Ratio
A	237	35.25	12.25	.67	.284
B	220	34.58	12.46		
A	237	34.25	12.25	.00	.002
C	219	34.25	12.99		
A	237	34.25	12.25	.48	.408
D	223	33.77	13.06		
B	220	34.58	12.46	.33	.272
C	219	34.25	12.99		
B	220	34.58	12.46	.81	.669
D	223	33.77	13.06		
C	219	34.25	12.99	.48	.387
D	223	33.77	13.06		

TABLE IX

Comparison of Mean Scores of Total Treatment Groups on the Metropolitan Matching Subtest

Group	N	Mean	S.D.	Diff. of Means	Critical Ratio
A	237	8.42	3.66	.04	.134
B	220	8.46	3.66		
A	237	8.42	3.66	.14	.438
C	219	8.56	3.36		
A	237	8.42	3.66	.37	1.112
D	223	8.05	3.44		
B	220	8.46	3.66	.10	.292
C	219	8.56	3.36		
B	220	8.46	3.66	.41	1.226
D	223	8.05	3.44		
C	219	8.56	3.36	.51	1.583
D	223	8.05	3.44		

Inspection of Tables VI-IX indicates that real differences did exist among the groups at the beginning of the experiment in spite of efforts to control such differences. They are as follows:

In intelligence as measured by the Pintner-Cunningham Primary Test: Mean scores fell in this order from low to high:

Group D - 33.51

Group A - 35.30 (control)

Group B - 35.99

Group C - 36.88

Differences between groups B and D and between groups C and D are significant at the 1% level. The difference between groups A and D is significant at the 5% level.

In knowledge of phonemes as measured by the Murphy-Durrell Phonemes Subtest: The mean score of group C differed from (was higher than) that of groups A, B, and D at the 1% level.

In knowledge of letter names as measured by the Murphy-Durrell Letter Names Subtest: There were no significant differences among the groups.

In the Metropolitan Matching Subtest: There were no significant differences among the groups.

Another important factor in comparing results of the various treatments is the attendance record of the children. Tables X and XI show the percentage of absence for each treatment group.

TABLE X

Percentage of School Absence during the Experimental Period of Children in Each Total Treatment Group

Group	N	Percentage of Absence
A	237	7.7
B	220	7.1
C	220	8.6
D	222	8.1

TABLE XI

Percentage of School Absence during the Experimental Period of Children in Low Subgroups within Total Treatment Groups

Group	N	Percentage of Absence
A	89	8.5
B	84	6.9
C	77	9.4
D	83	8.4

Teaching Staff

Teachers were chosen for the study; they were not volunteers. An effort was made to include teachers of all levels of experience and competence. Competence was initially rated by the principals of schools in which these teachers had worked, except in cases of beginning teachers, for whom no competence rating could be secured. Table XII shows the number of years of experience and the competence rating given the teachers in each treatment group. Principals were asked for a "general rating as a first grade teach-

er with emphasis on ability to teach reading." The rating scale was 4, 3, 2, 1 with 4 as top rating. In Table XII a question mark is used for beginning teachers with no experience and therefore no competency rating.

TABLE XII

Experience and Competency Rating for Teachers in Each Treatment Group

	Number in Group A	Number in Group B	Number in Group C	Number in Group D
Experience				
0 years	2	2	2	2
1-4 years	3	3	4	3
5 or more	5	5	4	5
Competency Rating				
?	2	2	2	2
1	0	0	0	0
2	1	1	1	1
3	5	5	4	4
4	2	2	3	3

As a further check on the competency of teachers in the four treatment groups, supervisors rated the teachers in May. Ratings were on a four-point scale with 4 as the top rating. These supervisor ratings are summarized in Table XIII.

TABLE XIII

Mean Ratings of Teacher Competency Made by Supervisors in May

Group	Mean Rating
A	2.9
B	2.9
C	2.7
D	2.7

These ratings would seem to indicate a high degree of similarity in the teacher competency in the four treatment groups. All teachers were women; all were certified teachers under the State Department of Education in Massachusetts. One teacher (in group B) resigned during the experimental period. She was replaced by a teacher of approximately equal qualifications.

The attendance record of teachers during the experimental period was good. Table XIV presents the percentage of absence.

TABLE XIV

Total Days and Percentage of School Absence during the Experimental Period of Teachers in Each Treatment Group

Group	Total Days of Teacher Absence	Percentage of Absence
A	25	1.87
B	38	2.77
C	59	4.27
D	55	3.97

Preliminary Testing

As soon after the opening of school as possible, the entire population was given the following tests. Tests were administered by the classroom teach-

ers with members of the reading department acting as assistants and observers.

1. Pintner-Cunningham Primary Test, Form A, Harcourt Brace and World, 1964. This is a standardized test of intelligence.
2. Murphy-Durrell Reading Readiness Analysis, Harcourt, Brace and World, 1964. This readiness test has three subtests:
   - Phonemes - a test of pupils' ability to identify separate sounds in spoken words
   - Letter Names - a test that requires identification of capital and lower case letters named by the examiner
   - Learning Rate - a test to determine the number of words a child is able to learn in one day under standardized conditions of presentation
3. Metropolitan Readiness Test, Form A, Harcourt, Brace and World, 1964. These subtests were given:
   - Word Meaning - a measure of pupils' store of verbal context
   - Listening - a test that is designed to tap the pupils' ability to comprehend phrases and sentences instead of individual words
   - Matching - a test of visual perception involving the recognition of similarities
4. Thurstone and Jeffrey Identical Forms Test, unpublished, a test of visual matching
5. Thurstone Pattern Copying Test, unpublished - a test in which the children complete each second figure to make it resemble the first
6. Letter Writing Test, unpublished - a test in which the children wrote letters named by the examiner

Raw scores from all of these tests were tabulated by classrooms for each pupil in the study. From these tabulations, about one-third of the children were chosen to become members of the "low subgroups" in their classrooms. These were children whose total testing profiles were lowest among the children in their classrooms. No citywide cutoff points could be established because of the wide variations in performance of children in different schools. The lowest scores obtained in certain classrooms in favored neighborhoods were sometimes found to be as high, or almost as high, as the highest scores obtained in certain other classrooms in disadvantaged neighborhoods. Therefore, each classroom had to be studied separately and those children selected who gave evidence in their test scores of being relatively least able within their own classrooms.

Teaching Procedures

Treatment A. The ten classrooms in Treatment A were used as controls. All children in these classes were taught with the regular basal program which had been used in their schools for several years prior to the experimental period: Scott Foresman (50's edition) or Ginn (1961 edition). Every effort was made to see that their program was "normal" in all respects. The "low subgroups" simply took the program more slowly than the more able groups. Teachers were asked to follow the manuals and to introduce nothing unusual into their teaching.

Treatment B. Ten classrooms used the same materials and procedures as in Treatment A. However, children in the "low subgroup" in each of these classrooms were given three half-hour additional teaching periods each week. This additional direct pupil-teacher contact was provided by two teachers specially trained in remedial reading who traveled from school to school to do this type of teaching. The two teachers exchanged schools midway in the experimental period. The traveling teachers worked on the

same word and other skills which the classroom teachers of these groups were working on, and there was a close team-teacher relationship between the classroom teachers and the traveling teachers. Time for this extra pupil-teacher contact was taken from the pupils' independent reading activities time. Their total time for reading was the same as that of other children in their rooms and in the total study.

Treatment C. Children in the "low subgroups" in these ten classrooms used different materials. Other children in these classrooms worked with their regular basal materials (Scott, Foresman or Ginn). The "low subgroups" were given intensive and prolonged training with the Houghton Mifflin readiness materials (Getting Ready to Read with its accompanying teaching devices: the basic card set, the plastic objects and boxes, and the Letto cards). When the children had achieved a firm mastery of the context-first consonant attack on words, which is the essence of the Houghton Mifflin procedure, they were introduced to a series of trade books, of which they read as many as time permitted. Trade books were used in an attempt to give children materials of greater intrinsic interest than basal readers so that they might put a greater amount of energy into working with them. The trade books were used in place of basal materials in group instruction in the "low subgroups." Teachers were trained to apply the context-first consonant approach to words introduced in the trade books (See Appendix A for a list of trade books used and for sample teaching materials given the teachers to assist them in using the trade books.)

Treatment D. The ten classrooms in Treatment D were given a combination of the procedures described under Treatments B and C. In these classrooms, as in Treatment C, children not assigned to the "low subgroups" were given the regular basal program of their school.

The time factor

The school year 1964-65 was 180 days. The experimental period of 140 days extended from October 21 to May 28. Prior to the beginning of the experimental period, the preliminary testing was done and results studied in order to determine the membership of the "low subgroups" within each classroom. While this study was going on, teachers were asked not to start any formal teaching of reading. They worked on informal readiness activities and made many experience charts with the children.

The school day in the elementary schools of Springfield is a two-session day: from 8:45 to 11:35 and from 1:00 to 3:15 except on Tuesday, when the afternoon session is from 1:00 to 2:30. The school week for first grade children is 24 hours.

The Springfield program of studies sets forth weekly time allotments for each area of the curriculum. In first grade 425 minutes per week are assigned to the teaching of reading and phonics. An additional 75 minutes per week are assigned to language development, mainly oral.

Teachers in all four treatment groups were asked to adhere to this standard time allotment, which refers to pupil, not teacher, time. The "reading and phonics" time allotment covered the following kinds of activities: group instruction by the teacher in whatever materials were assigned to the treatment group and independent seatwork activities related to the group instruction. The "language" time allotment covered such activities as oral language, reading to children, story telling, etc.

A check was taken toward the end of the year of teachers' adherence to this time allotment. This check was done by means of an individual interview with each teacher. The check revealed some variations. It is impossible to

tell whether these variations were significant enough to affect results. Probably the time factor is the most difficult to control in a study of this type.

Supervision

One person, Mrs. Mildred Lowe, was assigned full time to the supervision of the 40 classrooms. She found that Treatments C and D, in which teachers were using materials with which they were completely unfamiliar, needed the most help. More demonstration lessons and materials were prepared for teachers in these groups than for those in Treatment groups A and B. In treatments A and B, the teachers had the help of the manuals accompanying the basal materials, manuals with which most of them were already very familiar. Every effort was made to see that teachers in treatments A and B were doing the best job of which they were capable. Classroom observations, demonstration, professional meetings, and suggestions were provided for these teachers as well as for those in groups C and D.

ANALYSIS OF DATA

The first testing of results occurred at the end of January when two subtests of the Murphy-Durrell Reading Readiness Analysis (phonemes and letter names) and three subtests of the Metropolitan Readiness Test (word meaning, listening and matching) were readministered to children in the "low subgroups" in all treatment groups. Tables XV - XIX present the gains made between the two testing periods.

TABLE XV

Comparison of Mean Scores of Subgroups on the Murphy-Durrell Phonemes Subtest Administered in September and January

Subgroup	N	Means	S.D.	Diff. of Means	C.R.
A	86	Sept. 14.67 Jan. 29.56	Sept. 11.58 Jan. 14.19	14.89	7.56
B	81	Sept. 12.99 Jan. 31.75	Sept. 10.92 Jan. 12.97	18.76	9.97
C	74	Sept. 17.74 Jan. 32.20	Sept. 11.95 Jan. 15.16	14.46	6.45
D	81	Sept. 11.75 Jan. 28.11	Sept. 8.38 Jan. 14.50	16.36	8.79

TABLE XVI

Comparison of Mean Scores of Subgroups on the Murphy-Durrell Letter Names Subtest Administered in September and January

Subgroup	N	Means		S.D.		Diff. of Means	C.R.
A	90	Sept.	24.44	Sept.	10.54	17.34	11.48
		Jan.	41.78	Jan.	9.77		
B	81	Sept.	27.43	Sept.	11.94	17.22	10.56
		Jan.	44.65	Jan.	8.53		
C	75	Sept.	24.87	Sept.	11.02	19.93	11.93
		Jan.	44.80	Jan.	9.42		
D	81	Sept.	23.54	Sept.	10.64	21.05	13.76
		Jan.	44.59	Jan.	9.81		

TABLE XVII

Comparison of Mean Scores of Subgroups on the Metropolitan Word Meaning Subtest Administered in September and January

Subgroup	N	Means		S.D.		Diff. of Means	C.R.
A	90	Sept.	7.83	Sept.	3.55	.44	.89
		Jan.	8.27	Jan.	3.00		
B	81	Sept.	7.74	Sept.	2.64	1.07	2.43
		Jan.	8.81	Jan.	2.96		
C	76	Sept.	8.30	Sept.	2.81	1.27	2.82
		Jan.	9.57	Jan.	2.78		
D	82	Sept.	7.40	Sept.	2.80	1.03	2.14
		Jan.	8.43	Jan.	3.25		

TABLE XVIII

Comparison of Mean Scores of Subgroups on the Metropolitan Listening Subtest Administered in September and January

Subgroup	N	Means	S.D.	Diff. of Means	C. R.
A	90	Sept. 7.77 Jan. 8.93	Sept. 2.84 Jan. 2.29	1.16	3.05
B	81	Sept. 7.78 Jan. 9.22	Sept. 2.14 Jan. 2.34	1.44	4.00
C	76	Sept. 7.36 Jan. 9.37	Sept. 2.75 Jan. 2.24	2.01	4.90
D	82	Sept. 7.70 Jan. 8.94	Sept. 2.78 Jan. 2.45	1.24	3.02

TABLE XIX

Comparison of Mean Scores of Subgroups on the Metropolitan Matching Subtest Administered in September and January

Subgroup	N	Means	S.D.	Diff. of Means	C. R.
A	90	Sept. 6.53 Jan. 8.50	Sept. 3.43 Jan. 3.50	1.97	3.79
B	81	Sept. 6.37 Jan. 8.70	Sept. 3.93 Jan. 3.64	2.33	3.78
C	76	Sept. 6.92 Jan. 9.05	Sept. 3.29 Jan. 3.92	2.13	3.67
D	82	Sept. 6.09 Jan. 7.77	Sept. 3.52 Jan. 3.96	1.68	2.89

Inspection of Tables XV - XIX reveals very significant growth by all groups on the Murphy-Durrell Phonemes Subtest and the Murphy-Durrell Letter Names Subtest. Growth on the Metropolitan Word Meaning Subtest is significant at the 1% level only for Group C. On the Metropolitan Listening and Matching Subtests growth is significant at the 1% level for all groups. It appears that the type of program in the "low subgroups" had made little difference in the pattern of growth from group to group at this point in the experiment.

Analysis of June Data

At the end of the experimental period, the following group tests were given to all pupils in the four treatment groups:

Stanford Achievement Test, Primary 1, Form X

Test 1, Word Reading - a test consisting of 35 items which measures the child's ability to analyze a word with the aid of picture context. It employs a multiple-choice type of item in which the pupil looks at a picture and then selects one word which stands for that picture out of a group of four words given.

Test 2, Paragraph Meaning - a test of 38 items which places emphasis on comprehension of the material read.

Test 3, Vocabulary - a test which measures the child's vocabulary independent of his reading skill. Both questions and answers are read by the examiner. This test employs a multiple-choice type of item.

Test 4, Spelling - a test that employs a dictation-type exercise in which the word to be spelled is pronounced, illustrated in a sentence, and then written by the children.

Test 5, Word Study Skills - a test which measures phonetic skill:

initial sounds, final sounds, total sound of a word, and rhyming words.

Test 6, Arithmetic - a test that measures basic knowledge of standard units and number concepts, and evaluates the child's ability to do simple computation and understand the language of problems.

San Diego Attitude Inventory, Department of Education, San Diego County - a test, consisting of 25 questions, that measures the pupils' feeling about reading. Questions are read by the examiner and the pupils respond by circling Yes or No. Sample questions are: Do you like to read before you go to bed? Do you think you are a poor reader?

The following tests were administered individually to a random sampling of pupils in all four treatment groups. The administration of the tests to the random sample was done by members of the study staff and by trained reading teachers.

Gilmore Oral Reading Test, Form A. In the administration of this test, only two scores were recorded: accuracy (a test of pupils' ability to pronounce words in context) and rate. Hesitations and repetitions were not counted as errors. Hence, norms published for this test are not appropriate in this study. Types of errors checked to obtain the accuracy score were: substitutions, mispronunciations, words pronounced by examiner, disregard of punctuations, insertions, and omissions.

Fry Word Pronunciation Test - a test of pupils' ability to pronounce out of context a list of phonetically regular words.

Karlsen Phonemic Word Test - a measure primarily of ability to apply phonetic principles

Gates Word Pronunciation Test - a test in which children are asked to pronounce a list of graded words out of context

The children also took two tests of writing ability:

Writing Sample 1. Restricted Stimulus - Pupils were encouraged to write a story with motivation of a very general type. They were directed to write on anything they chose for twenty minutes without help.

Writing Sample 2, Unique Stimulus - a test that gave the pupils an opportunity to write stories using a definite motivational stimulus, whatever type of stimulus they were accustomed to.

Tables XX through XXV record the data from the posttests for the total treatment groups. All mean scores are raw scores.

TABLE XX

Comparison of Mean Scores of Total Treatment Groups on Stanford Word Reading Subtest Administered in June

Group	N	Mean	S.D.	Diff. of Means	C.R.
A	237	18.73	6.10	.68	1.253
B	220	19.41	5.63		
A	237	18.73	6.10	1.08	1.750
C	221	19.81	7.08		
A	237	18.73	6.10	1.26	2.085
D	221	19.99	6.83		
B	220	19.41	5.63	.40	.651
C	221	19.81	7.08		
B	220	19.41	5.63	.58	.968
D	221	19.99	6.83		
C	221	19.81	7.08	.18	.273
D	221	19.99	6.83		

TABLE XXI

Comparison of Mean Scores of Total Treatment Groups on Stanford Paragraph Meaning Subtest Administered in June

Group	N	Mean	S.D.	Diff. of Means	C.R.
A	237	19.71	8.07	.04	.059
B	220	19.67	8.23		
A	237	19.71	8.07	1.06	1.257
C	221	20.77	9.83		
A	237	19.71	8.07	1.15	1.301
D	221	20.86	10.47		
B	220	19.67	8.23	1.10	1.281
C	221	20.77	9.83		
B	220	19.67	8.23	1.19	1.324
D	221	20.86	10.47		
C	221	20.77	9.83	.09	.084
D	221	20.86	10.47		

TABLE XXII

Comparison of Mean Scores of Total Treatment Groups on Stanford Vocabulary Subtest Administered in June

Group	N	Mean	S.D.	Diff. of Means	C.R.
A	237	21.02	6.20	.74	1.289
B	220	21.76	6.17		
A	237	21.02	6.20	1.03	1.789
C	221	22.05	6.14		
A	237	21.02	6.20	1.05	1.787
D	221	19.97	6.29		
B	220	21.76	6.17	.29	.488
C	221	22.05	6.14		
B	220	21.76	6.17	1.79	3.018
D	221	19.97	6.29		
C	221	22.05	6.14	2.08	3.512
D	221	19.97	6.29		

TABLE XXIII

Comparison of Mean Scores of Total Treatment Groups on Stanford Spelling Subtest Administered in June

Group	N	Mean	S.D.	Diff. of Means	C.R.
A	237	11.51	5.55	.43	.828
B	220	11.94	5.44		
A	237	11.51	5.55	.36	.611
C	221	11.15	6.96		
A	237	11.51	5.55	.76	1.379
D	221	10.75	6.25		
B	220	11.94	5.44	.79	1.323
C	221	11.15	6.96		
B	220	11.94	5.44	1.19	2.132
D	221	10.75	6.25		
C	221	11.15	6.96	.40	.640
D	221	10.75	6.25		

TABLE XXIV

Comparison of Mean Scores of Total Treatment Groups on Stanford Word Study Skills Subtest Administered in June

Group	N	Mean	S.D.	Diff. of Means	C.R.
A	237	36.03	8.93	.78	.997
B	220	36.81	7.89		
A	237	36.03	8.93	1.27	1.507
C	221	37.30	9.08		
A	237	36.03	8.93	1.34	1.493
D	221	37.37	10.20		
B	220	36.81	7.39	.49	.599
C	221	37.30	9.08		
B	220	36.81	7.89	.56	.642
D	221	37.37	10.20		
C	221	37.30	9.08	.07	.079
D	221	37.37	10.20		

The Stanford Subtests revealed the following significant differences among the four treatment groups:

- in the Word Reading Subtest Group D was superior to Group A (control) at the 5% level.

- in the Vocabulary Subtest, Group B was superior to Group D at the 1% level. Group C was also superior to Group D at the 1% level.

- in the Spelling Subtest Group B was superior to Group D at the 5% level. Other differences, while generally in favor of the experimental groups over the control, were not significant.

TABLE XXV

Comparison of Mean Scores of Total Treatment Groups on the San Diego Attitude Inventory Administered in June

Group	N	Mean	S.D.	Diff. of Means	C.R.
A	237	16.73	4.36	.47	1,088
B	220	17.20	4.84		
A	237	16.73	4.36	.82	1.837
C	221	17.55	5.06		
A	237	16.73	4.36	.69	1.684
D	221	17.42	4.30		
B	220	17.20	4.84	.35	.727
C	221	17.55	5.06		
B	220	17.20	4.84	.22	.485
D	221	17.42	4.30		
C	221	17.55	5.06	.13	.294
D	221	17.42	4.30		

Although scores on the Attitude Inventory were generally higher for experimental groups than for the control group, there were no statistically significant differences.

Because this study was focussed particularly upon the relative performance of children in the "low subgroups" within each treatment group, separate studies were made of test results for these children. Tables XXVI- XXX summarize the results for the subtests of the Stanford Test.

TABLE XXVI

Comparison of Mean Scores of Low Subgroups within the Total Treatment Groups on the Stanford Word Reading Test

Subgroup	N	Mean	S.D.	Diff. of Means	C. R.
A	79	15.99	4.37	.15	.239
B	81	16.14	3.42		
A	79	15.99	4.37	.95	1.324
C	64	16.94	4.19		
A	79	15.99	4.37	.98	1.203
D	69	16.97	5.43		
B	81	16.14	3.42	.80	1.240
C	64	16.94	4.19		
B	81	16.14	3.42	.83	1.105
D	69	16.97	5.43		
C	64	16.94	4.19	.03	.040
D	69	16.97	5.43		

TABLE XXVII

Comparison of Mean Scores of Low Subgroups within the Total Treatment Groups on Stanford Paragraph Meaning Subtest

Subgroup	N	Mean	S.D.	Diff. of Means	C.R.
A	79	14.73	5.29	.38	.468
B	81	14.35	5.20		
A	79	14.73	5.29	1.06	1.087
C	64	13.67	6.20		
A	79	14.73	5.29	.53	.365
D	69	15.26	10.93		
B	81	14.35	5.20	.68	.698
C	64	13.67	6.20		
B	81	14.35	5.20	.91	.637
D	69	15.26	10.93		
C	64	13.67	6.20	1.59	1.041
D	69	15.26	10.93		

TABLE XXVIII

Comparison of Mean Scores of Low Subgroups within the Total Treatment Groups on Stanford Vocabulary Subtest

Subgroup	N	Mean	S.D.	Diff. of Means	C.R.
A	79	18.37	5.66	.27	.316
B	81	18.64	5.34		
A	79	18.37	5.66	1.11	1.201
C	64	19.48	5.42		
A	79	18.37	5.66	2.21	2.571
D	69	16.16	4.78		
B	81	18.64	5.34	.84	.936
C	64	19.48	5.42		
B	81	18.64	5.34	2.48	3.005
D	69	16.16	4.78		
C	64	19.48	5.42	3.32	3.741
D	69	16.16	4.78		

TABLE XXIX

Comparison of Mean Scores of Low Subgroups within the Total Treatment Groups on the Stanford Spelling Subtest

Subgroup	N	Mean	S.D.	Diff. of Means	C.R.
A	79	8.92	5.06	.13	.167
B	81	9.05	4.38		
A	79	8.92	5.06	2.17	2.584
C	64	6.75	4.96		
A	79	8.92	5.06	1.25	1.471
D	69	7.67	5.30		
B	81	9.05	4.38	2.30	2.918
C	64	6.75	4.96		
B	81	9.05	4.38	1.38	1.724
D	69	7.67	5.30		
C	64	6.75	4.96	.92	1.031
D	69	7.67	5.30		

TABLE XXX

Comparison of Mean Scores of Low Subgroups within the Total Treatment Groups on Stanford Word Study Skills Subtest

Subgroup	N	Mean	S.D.	Diff. of Means	C.R.
A	79	30.59	6.87	2.20	2.110
B	81	32.79	6.27		
A	79	30.59	6.87	2.36	2.150
C	64	32.95	6.22		
A	79	30.59	6.87	3.06	2.148
D	69	33.65	9.93		
B	81	32.79	6.27	.16	.156
C	64	32.95	6.22		
B	81	32.79	6.27	.86	.623
D	69	33.65	9.93		
C	64	32.95	6.22	.70	.490
D	69	33.65	9.93		

Inspection of Tables XXVI-XXX indicates the following:

Stanford Word Reading Subtest - All experimental groups were superior to the control group but none significantly.

Stanford Paragraph Meaning Subtest - The control group (A) was superior to groups B and C but not significantly. Group D was superior to the control group but not significantly.

Stanford Vocabulary Subtest - The control group (A) was superior to group D at the 5% level. Groups B and C were superior to group D at the 1% level. Groups B and C were superior to the control group but neither significantly.

Stanford Spelling Subtest - The control group (A) was superior to group C at the 1% level and superior to group D but not significantly. Group B was superior to the control group but not significantly. Group B was superior to group C at the 1% level.

Stanford Word Study Subtest - Groups B, C, and D were all superior to the control group at the 5% level.

Results of the San Diego Attitude Inventory in the low subgroups is summarized in Table XXXI

TABLE XXXI

Comparison of Mean Scores of Subgroups within the Total Treatment Groups on the San Diego Attitude Inventory

Subgroup	N	Mean	S.D.	Diff. of Means	C.R.
A	79	15.96	3.97	1.35	2.029
B	81	17.31	4.42		
A	79	15.96	3.97	1.32	1.730
C	64	17.28	4.95		
A	79	15.96	3.97	1.66	2.509
D	69	17.62	4.06		
B	81	17.31	4.42	.03	.035
C	64	17.28	4.95		
B	81	17.31	4.42	.31	.454
D	69	17.62	4.06		
C	64	17.28	4.95	.34	.434
D	69	17.62	4.06		

Table XXXI shows that groups B and D were superior to the control group (A) at the 5% level. Group C was also superior to the control group but not significantly. There was little difference among the three experimental groups.

Testing of a random sampling of each of the total treatment groups is summarized in TABLES XXXII - XLI

TABLE XXXII

Comparison of Mean Scores of Random Sample within Total Treatment Groups on Gilmore Accuracy Test Administered in June

Group	N	Mean	S.D.	Diff. of Means	C.R.
A	45	23.13	9.65	.78	.442
B	52	22.35	7.56		
A	45	23.13	9.65	4.04	1.824
C	48	27.17	11.64		
A	45	23.13	9.65	1.69	.780
D	44	24.82	10.69		
B	52	22.35	7.56	4.82	2.435
C	48	27.17	11.64		
B	52	22.35	7.56	2.47	1.285
D	44	24.82	10.69		
C	48	27.17	11.64	2.35	1.009
D	44	24.82	10.69		

TABLE XXXIII

Comparison of Mean Scores of Random Sample within Total Treatment Groups on Gilmore Rate Subtest Administered in June

Group	N	Mean	S. D.	Diff. of Means	C. R.
A	45	50.76	17.02	.97	.294
B	52	51.73	15.46		
A	45	50.76	17.02	3.72	.877
C	48	54.48	23.57		
A	45	50.76	17.02	6.31	1.528
D	44	57.07	21.64		
B	52	51.73	15.46	2.75	.683
C	48	54.48	23.57		
B	52	51.73	15.46	5.34	1.367
D	44	57.07	21.64		
C	48	54.48	23.57	2.59	.549
D	44	57.07	21.64		

TABLE XXXIV

Comparison of Mean Scores of Random Sample within Total Treatment Groups on Fry Word Pronunciation Test Administered in June

Group	N	Mean	S. D.	Diff. of Means	C. R.
A	33	7.36	6.61	1.62	1.152
B	35	5.74	4.79		
A	33	7.36	6.61	2.27	1.365
C	40	9.63	7.54		
A	33	7.36	6.61	3.41	2.000
D	30	10.77	6.87		
B	35	5.74	4.79	3.89	2.695
C	40	9.63	7.54		
B	35	5.74	4.79	5.03	3.367
D	30	10.77	6.87		
C	40	9.63	7.54	1.14	.660
D	30	10.77	6.87		

TABLE XXXV

Comparison of Mean Scores of Random Sample within Total Treatment Groups on the Karlsen Phonemic Word Test Administered in June

Group	N	Mean	S. D.	Diff. of Means	C. R.
A	45	9.00	7.88	1.47	1.031
B	51	7.53	5.77		
A	45	9.00	7.88	2.17	1.224
C	48	11.17	9.18		
A	45	9.00	7.88	2.18	1.251
D	44	11.18	8.55		
B	51	7.53	5.77	3.64	2.345
C	48	11.17	9.18		
B	51	7.53	5.77	3.65	2.400
D	44	11.18	8.55		
C	48	11.17	9.18	.01	.008
D	44	11.18	8.55		

TABLE XXXVI

Comparison of Mean Scores of Random Sample Within the Total Treatment Groups on Gates Word Pronunciation Test Administered in June

Group	N	Mean	S. D.	Diff. of Means	C. R.
A	45	11.67	6.22	.96	.861
B	52	10.71	4.39		
A	45	11.67	6.22	1.83	1.328
C	48	13.50	7.09		
A	45	11.67	6.22	1.33	.984
D	44	13.00	6.56		
B	52	10.71	4.39	2.79	2.342
C	48	13.50	7.09		
B	52	10.71	4.39	2.29	1.971
D	44	13.00	6.56		
C	48	13.50	7.09	.50	.352
D	44	13.00	6.56		

TABLE XXXVII

Comparison of Mean Scores of Random Sample Within the Total Treatment Groups on Writing Sample: Mechanics Ratio Scale (Restricted Stimulus)

Group	N	Mean	S. D.	Diff. of Means	C. R.
A	41	60.93	20.37	1.41	.318
B	44	59.52	20.36		
A	41	60.93	20.37	6.22	.901
C	41	67.15	39.24		
A	41	60.93	20.37	.35	.073
D	38	60.58	21.88		
B	44	59.52	20.36	7.63	1.112
C	41	67.15	39.24		
B	44	59.52	20.36	1.06	.225
D	38	60.58	21.88		
C	41	67.15	39.24	6.57	.927
D	38	60.58	21.88		

TABLE XXXVIII

Comparison of Mean Scores of Random Sample Within the Total Treatment Groups on Writing Sample: Total Number of Words Spelled Correctly (Restricted Stimulus)

Group	N	Mean	S. D.	Diff. of Means	C. R.
A	42	27.40	11.85	2.16	.718
B	46	25.24	16.27		
A	42	27.40	11.85	3.69	1.391
C	42	23.71	12.47		
A	42	27.40	11.85	.11	.034
D	39	27.51	15.99		
B	46	25.24	16.27	1.53	.496
C	42	23.71	12.47		
B	46	25.24	16.27	2.27	.648
D	39	27.51	15.99		
C	42	23.71	12.47	3.80	1.186
D	39	27.51	15.99		

TABLE XXXIX

Comparison of Mean Scores of Random Sample Within the Total Treatment Groups on Writing Sample: Total Number of Running Words (Restricted Stimulus)

Group	N	Mean	S. D.	Diff. of Means	C. R.
A	42	32.36	13.78	2.82	.865
B	46	29.54	16.68		
A	42	32.36	13.78	3.96	1.276
C	42	28.40	14.59		
A	42	32.36	13.78	.00	.001
D	39	32.36	16.49		
B	46	29.54	16.68	1.14	.342
C	42	28.40	14.59		
B	46	29.54	16.68	2.82	.780
D	39	32.36	16.49		
C	42	28.40	14.59	3.96	1.140
D	39	32.36	16.49		

TABLE XL

Comparison of Mean Scores of Random Sample within Total Treatment Groups on Writing Sample: Total Number of Words Spelled Correctly (Unique Stimulus)

Group	N	Mean	S.D.	Diff. of Means	C.R.
A	41	24.73	12.48	7.22	2.290
B	44	31.95	16.45		
A	41	24.73	12.48	2.29	.731
C	43	27.02	16.11		
A	41	24.73	12.48	6.57	1.866
D	44	31.30	19.43		
B	44	31.95	16.45	4.93	1.413
C	43	27.02	16.11		
B	44	31.95	16.45	.65	.172
D	44	31.30	19.43		
C	43	27.02	16.11	4.28	1.118
D	44	31.30	19.43		

TABLE XLI

Comparison of Mean Scores of Random Sample within Total Treatment Groups on Writing Sample: Total Number of Running Words (Unique Stimulus)

Group	N	Mean	S.D.	Diff. of Means	C.R.
A	41	27.66	13.17	9.84	2.896
B	44	37.50	17.94		
A	41	27.66	13.17	4.50	1.301
C	43	32.16	18.25		
A	41	27.66	13.17	7.52	2.014
D	44	35.18	20.69		
B	44	37.50	17.94	5.34	1.375
C	43	32.16	18.25		
B	44	37.50	17.94	2.32	.562
D	44	35.18	20.69		
C	43	32.16	18.25	3.02	.722
D	44	35.18	20.69		

Tables XXXII-XLI reveal the following significant differences:

Gilmore Accuracy Test

Group C superior to Group B at the 5% level.

Gilmore Rate Test

No significant differences among groups.

Fry Word Pronunciation Test

Group D superior to Group A at the 5% level.

Group C superior to Group B at the 1% level.

Group D superior to Group B at the 1% level.

Karlsen Phonemic Word Test

Group C superior to Group B at the 5% level.

Group D superior to Group B at the 5% level.

Gates Word Pronunciation Test

Group C superior to Group B at the 5% level.

Group D superior to Group B at the 5% level.

Restricted Stimulus Writing Sample: Mechanics Ratio Scale

No significant differences among groups.

Restricted Stimulus Writing Sample: Number of Words Spelled Correctly

No significant differences among groups.

Unique Stimulus Writing Sample: Total Number of Running Words

No significant differences among groups.

Unique Stimulus Writing Sample: Number of Words Spelled Correctly

Group B superior to Group A at the 5% level.

Unique Stimulus Writing Sample: Total Number of Running Words

Group B superior to Group A at the 1% level.

Group D superior to Group A at the 5% level.

Again, because the emphasis in the study was on the "slow subgroups" it seemed important to compare the results on the individually administered tests of children in these low subgroups. Because most of these tests were given only to a random sample of the total groups, the number of cases which fell within each of the low subgroups is small. The exceptions are the Fry Word Pronunciation Test and the Gates Word Pronunciation Test which were given to a random sample of the total population but to all children in the low subgroups. Tables XLII-XLIX summarize the results.

TABLE XLII

Comparison of Mean Scores of Children from Random Sampling Who Fell Within the Low Subgroups on the Gilmore Accuracy Subtest

Group	N	Means	S. D.	Diff. of Means	C.R.
A	16	15.00	4.61	1.72	.952
B	18	16.72	5.61		
A	16	15.00	4.61	5.07	2.007
C	14	20.07	8.02		
A	16	15.00	4.61	1.07	.414
D	14	16.07	8.27		
B	18	16.72	5.61	3.35	1.283
C	14	20.07	8.02		
B	18	16.72	5.61	.65	.243
D	14	16.07	8.27		
C	14	20.07	8.02	4.00	1.249
D	14	16.07	8.27		

TABLE XLIII

Comparison of Mean Scores of Children from Random Sampling Who Fell Within the Low Subgroups on Gilmore Rate Subtest

Group	N	Means	S. D.	Diff. of Means	C.R.
A	16	37.13	14.73	11.15	2.130
B	18	48.28	14.79		
A	16	37.13	14.73	.58	.117
C	14	37.71	11.79		
A	16	37.13	14.73	3.58	.591
D	14	40.71	17.08		
B	18	48.28	14.79	10.57	2.175
C	14	37.71	11.79		
B	18	48.28	14.79	7.57	1.274
D	14	40.71	17.08		
C	14	37.71	11.79	3.00	.522
D	14	40.71	17.08		

TABLE XLIV

Comparison of Mean Scores of Children in the Low Subgroups on the Fry Word Pronunciation Test

Group	N	Means	S.D.	Diff. of Means	C.R.
A	86	.74	1.66	.52	1.575
B	81	1.26	2.47		
A	86	.74	1.66	1.55	3.441
C	76	2.29	3.60		
A	86	.74	1.66	1.67	3.280
D	81	2.41	4.27		
B	81	1.26	2.47	1.03	2.081
C	76	2.29	3.60		
B	81	1.26	2.47	1.15	2.095
D	81	2.41	4.27		
C	76	2.29	3.60	.12	.188
D	81	2.41	4.27		

TABLE XLV

Comparison of Mean Scores of Children in the Low Subgroups on the Gates Word Pronunciation Test

Group	N	Means	S.D.	Diff. of Means	C.R.
A	86	6.69	3.05	.46	.960
B	82	7.15	3.15		
A	86	6.69	3.05	.52	.939
C	76	7.21	3.95		
A	86	6.69	3.05	1.50	2.350
D	81	8.19	4.92		
B	82	7.15	3.15	.06	.114
C	76	7.21	3.95		
B	82	7.15	3.15	1.04	1.603
D	81	8.19	4.92		
C	76	7.21	3.95	.98	1.372
D	81	8.19	4.92		

TABLE XLVI

Comparison of Mean Scores of Children from Random Sampling Who Fell Within the Low Subgroups on the Karlsen Phonemic Word Test

Group	N	Means	S.D.	Diff. of Means	C.R.
A	16	3.63	1.27	.07	.089
B	18	3.56	2.89		
A	16	3.63	1.27	1.87	1.717
C	14	5.50	3.76		
A	16	3.63	1.27	3.23	1.765
D	14	6.86	6.50		
B	18	3.56	2.89	1.94	1.549
C	14	5.50	3.76		
B	18	3.56	2.89	3.30	1.709
D	14	6.86	6.50		
C	14	5.50	3.76	1.36	.652
D	14	6.86	6.50		

TABLE XLVII

Comparison of Mean Scores of Children from Random Sampling Who Fell Within the Low Subgroups on Writing Sample: Mechanics Ratio Scale (Restricted Stimulus)

Group	N	Means	S.D.	Diff. of Means	C. R.
A	15	58.33	23.05	6.45	.779
B	17	51.88	22.17		
A	15	58.33	23.05	2.12	.222
C	14	56.21	26.28		
A	15	58.33	23.05	17.10	1.682
D	13	41.23	27.99		
B	17	51.88	22.17	4.33	.474
C	14	56.21	26.28		
B	17	51.88	22.17	10.65	1.086
D	13	41.23	27.99		
C	14	56.21	26.28	14.98	1.377
D	13	41.23	27.99		

TABLE XLVIII

Comparison of Mean Scores of Children from Random Sampling Who Fell Within the Low Subgroups on Writing Sample: Total Number of Words Spelled Correctly (Restricted Stimulus)

Group	N	Means	S. D.	Diff. of Means	C.R.
A	15	22.87	9.29	1.55	.280
B	17	24.41	19.71		
A	15	22.87	9.29	1.73	.443
C	14	21.14	10.82		
A	15	22.87	9.29	1.90	.371
D	13	24.77	15.54		
B	17	24.41	19.71	3.27	.566
C	14	21.14	10.82		
B	17	24.41	19.71	.36	.053
D	13	24.77	15.54		
C	14	21.14	10.82	3.63	.671
D	13	24.77	15.54		

TABLE XLIX

Comparison of Mean Scores of Children from Random Sampling Who Fell Within the Low Subgroups on Writing Sample: Total Number of Running Words (Restricted Stimulus)

Group	N	Means	S. D.	Diff. of Means	C. R.
A	15	27.27	10.00	.50	.089
B	17	27.77	19.71		
A	15	27.27	10.00	.02	.004
C	14	27.29	13.70		
A	15	27.27	10.00	3.50	.627
D	13	30.77	16.99		
B	17	27.77	19.71	.48	.077
C	14	27.29	13.70		
B	17	27.77	19.71	3.00	.432
D	13	30.77	16.99		
C	14	27.29	13.70	3.48	.561
D	13	30.77	16.99		

Inspection of Tables XLII-XLIX reveals the following:

Gilmore Accuracy Test

Group C superior to Group A at the 5% level.

Gilmore Rate Test

Group B superior to Group A at the 5% level.

Group B superior to Group C at the 5% level.

Fry Word Pronunciation Test

Group C superior to Group A at the 1% level.

Group D superior to Group A at the 1% level.

Group C superior to Group B at the 5% level.

Group D superior to Group B at the 5% level.

Gates Word Pronunciation Test

Group D superior to Group A at the 5% level.

Karlsen Phonemic Word Test

No significant differences among groups.

Writing Sample: Mechanics Ratio Scale (Restricted Stimulus)

No significant differences among groups.

Writing Sample: Total Number of Words Spelled Correctly (Restricted Stimulus)

No significant differences among groups.

Writing Sample: Total Number of Running Words (Restricted Stimulus)

No significant differences among groups.

Informal Observations

The following observations should be noted:

1. Children in the groups using the Houghton Mifflin materials and the trade books (the low subgroups in total groups C and D) showed a remarkable eagerness for reading and a confidence unusual with children in slow groups. We believe this was due to two factors: the basic skills developed in the longer and more intensive readiness work and the fact that these children had reading materials (trade books) different from those used by other children in their classrooms. The stories were not "worn out" before the slow children read them.
2. Strict vocabulary control seems to be unnecessary if children are really ready to attack words before they begin to read. The trade books posed no problem, and children in the low groups picked up preprimers and primers from the basal series other children in their rooms were using and read them at sight.
3. The amount of reading by some of the low groups in total groups C and D was considerably more than low groups usually accomplish even though they spent fewer days in reading (because of the extended readiness period). Some of the low groups read as many as ten trade books. Usually low groups are considered to have done well if they complete the primer in a basal series (usually a total of four books).

Predictive Value of the Pretests

Children most likely to have difficulty in learning to read were identified by means of the battery of pretests already described. Informally, the procedure seemed effective. No children had to be moved into the "low subgroups" later in the year because the tests had not identified them earlier. A very few children seemed to "blossom" during the year and they were moved into faster groups.

The pretesting procedures, though apparently effective, were long and clumsy in operation. Few schools would wish to administer such an array of tests on a permanent basis. The predictive value of each of the pretests was estimated by finding the Pearson product-moment correlations between scores on each pretest and the following selected posttests: Stanford subtests of Word Reading, Paragraph Meaning, and Word Study Skills, the Gilmore Accuracy Test, and the Fry and Gates Word Pronunciation Tests. These correlations are reported in Tables L and LI.

TABLE L

Pearson Product Moment Correlations between Pretests and Selected Achievement Tests Administered to Total Treatment Groups

	Stanford Word Reading	Stanford Paragraph Meaning	Stanford Word Study Skills
Pintner	.2795	.3250	.2757
Murphy-Durrell Phonemes	.2424	.2793	.2327
Murphy-Durrell Letter Names	.2506	.3139	.2815
Murphy-Durrell Learning Rate	.1837	.2363	.2065
Metropolitan Word Meaning	.1555	.2028	.1469
Metropolitan Listening	.1507	.1907	.1594
Metropolitan Matching	.2636	.2998	.2518
Thurstone Identical Forms	.1813	.2089	.1579
Thurstone Pattern Copying	.1415	.1484	.1287
Informal Letter Writing	.5023	.5022	.4875

TABLE LI

Pearson Product Moment Correlations between Pretests and Selected Achievement Tests Administered to Random Samples of the Total Treatment Groups

	Gilmore Accuracy	Fry	Gates
Pintner	.1337	.0518	.0905
Murphy-Durrell Phonemes	.1152	.0575	.1061
Murphy-Durrell Letter Names	.1312	.0287	.1374
Murphy-Durrell Learning Rate	.1087	.0744	.0401
Metropolitan Word Meaning	.1870	.0932	.0082
Metropolitan Listening	.0512	.0091	.0886
Metropolitan Matching	.1315	.0225	.0521
Thurstone Identical Forms	.0566	.0122	.1255
Thurstone Pattern Copying	.0178	.0112	.0607
Informal Letter Writing	.1195	.1491	.0287

Tables L and LI indicate that in this situation no single test of the many administered could have been used with any confidence as a predictor of achievement in June. Interestingly, the simple little homemade test of Informal Letter Writing has the highest correlation with four of the six achievement scores.

Tables LII and LIII record the correlations between the pretests given to children in the low subgroups and selected achievement tests.

TABLE LII

Pearson Product Moment Correlations between Pretests and Selected Achievement Tests Administered to Low Subgroups

	Stanford Word Meaning	Stanford Paragraph Meaning	Stanford Word Study Skills
Pintner	.2794	.2550	.3560
Murphy-Durrell Phonemes	.2032	.1007	.2749
Murphy-Durrell Letter Names	.3629	.2509	.3575
Murphy-Durrell Learning Rate	.2344	.2133	.2662
Metropolitan Word Meaning	.1803	.1401	.2114
Metropolitan Listening	.2608	.1412	.2240
Metropolitan Matching	.2724	.2436	.3229
Thurstone Identical Forms	.2011	.1821	.2165
Thurstone Pattern Copying	.2005	.1625	.2602
Informal Letter Writing	.3060	.2408	.2899

TABLE LIII

Pearson Product Moment Correlations Between Pretests and Selected Achievement Tests Administered to Children from the Random Sampling Who Fell Within the Low Subgroups

	Gilmore Accuracy	Fry	Gates
Pintner	.0016	.1261	.2240
Murphy-Durrell Phonemes	.0324	.1667	.2187
Murphy-Durrell Letter Names	.0675	.3155	.4360
Murphy-Durrell Learning Rate	.1100	.2680	.3018
Metropolitan Word Meaning	.1033	.1133	.1322
Metropolitan Listening	.0391	.0679	.1080
Metropolitan Matching	.0177	.0734	.1327
Thurstone Identical Forms	.0306	.1224	.1330
Thurstone Pattern Copying	.0562	.0496	.0804
Informal Letter Writing	.1159	.2307	.3164

With children in the low subgroups the Murphy-Durrell Letter Names subtest proved to have the highest correlation with the achievement tests as a group.

LIMITATIONS OF THE STUDY

1. Groups using the Houghton Mifflin readiness materials and the trade books (Groups C and D) may have been penalized by the teachers' complete lack of prior familiarity with these materials. Most teachers of groups using the regular basal program (Groups A and B) had had considerable experience with the basal materials.
2. The San Diego Attitude Inventory may have doubtful validity for measuring the kind of enthusiasm for reading which seemed to be generated in groups using the trade books (Groups C and D).
3. Some of the achievement tests, particularly those stressing recognition of words in isolation, may be invalid for measuring results of a program in which heavy emphasis was placed upon teaching children to use context clues in recognizing words.
4. Because of problems in negotiation of the contract, it was not known for certain in Springfield that the project could be carried out until the last week before school opened in September. It was, therefore, not possible to give the participating teachers the type and extent of orientation to the plan which might have made it possible to get off to a better start. Under the circumstances, the experimental period did not begin until October 21. During the weeks prior to this date, which were spent in testing, grouping, and making other plans, it was difficult to control the type of program teachers were carrying on with the children.

CONCLUSIONS AND IMPLICATIONS

Two null hypotheses were tested in the experiment:

1. There is no significant difference in the distribution of test scores at the end of grade one among the "low subgroups" within the total treatment groups in each of the four situations (one control and three experimental).
2. There is no significant difference in the distribution of test scores at the end of grade one among the total treatment groups in each of the four situations.

The evidence is not completely consistent; however, there is more evidence for rejecting the hypotheses than for retaining them.

Although the groups of children and teachers were not perfectly matched (see pp. 15-24), it seems likely that some of the differences may have balanced each other. For example, Group C had the highest mean intelligence and scores on the Murphy-Durrell Phonemes subtest significantly higher than any other group (1% level). However, Group C also had the highest pupil absence, the highest teacher absence, and a preponderance of boys over girls (122 to 98).

Group D seems to have started at something of a disadvantage: mean intelligence significantly lower (5% level) than the control group and lower than groups B and C (1% level); mean score on the Murphy-Durrell Phonemes subtest significantly lower than group C (1% level).

The control group started with a mean score on the Murphy-Durrell Phonemes subtest significantly lower than group C (1% level).

These facts must be kept in mind in evaluating results.

Because the large amount of evidence presented in the tables makes it difficult for the reader to see the trends, eight additional

tables have been constructed to summarize the most important data. In Summary Tables 1-4 an x indicates that the mean of the experimental group exceeds the mean of the control group. If the difference is significant, 5% or 1% is entered in the space to show the degree of significance.

SUMMARY TABLE 1

Comparison of Experimental Groups with Control Group (A) on Tests Dependent Mainly on Skill in Word Recognition (Low Subgroups)

Experimental Group	Stanford Word Reading	Stanford Word Study Skills	Gilmore Accuracy	Gilmore Rate	Fry	Gates	Karlsen
B	x	5%	x	5%	x	x	
C	x	5%	5%	x	1%	x	x
D	x	5%	x	x	1%	5%	x

In 20 of the 21 comparisons made the mean of the experimental low subgroup exceeds the mean of the control low subgroup. It is probably important that the most significant differences occurred in the Stanford Word Study Skills subtest, which is heavily dependent on ability to use sounds within a word, and in the Fry Test of Phonetically Regular Words, although the children had not worked on vowel principles which are important in the Fry test.

SUMMARY TABLE 2

Comparison of Experimental Groups with Control Group (A) on Tests Dependent Mainly on Skill in Word Recognition (Total Treatment Groups)

Experimental Group	Stanford Word Reading	Stanford Word Study Skills	Gilmore Accuracy	Gilmore Rate	Fry	Gates	Karlsen
B	x	x		x			
C	x	x	x	x	x	x	x
D	5%	x	x	x	5%	x	x

In 17 of the 21 comparisons made the mean of the total experimental group exceeds the mean of the total control group, though usually not significantly. It should be recalled that no special procedures or materials were used in the experimental groups <u>except with children in the low subgroups</u>. It would appear, however, that there was some effect upon the total class. It should be noted that while the differences in favor of groups C and D are consistent, they are not so in group B where additional teacher time with the low subgroup was the only experimental variable.

SUMMARY TABLE 3

Comparison of Experimental Groups with Control Group (A) on Tests <u>Not</u> Primarily Dependent on Word Recognition (Low Subgroups)

Experimental Group	Stanford Paragraph Meaning	Stanford Vocabulary (oral)	Stanford Spelling	San Diego Attitude
B		x	x	5%
C		x	1% (C)	x
D	x	5% (C)		5%

(C) indicates that control exceeds experimental.

Since the skills measured by the Stanford Paragraph Meaning, Vocabulary, and Spelling subtests were not directly emphasized in the experimental work, it is not surprising that no clear trend emerges. In attitude toward reading, all three experimental low subgroups scored higher on the San Diego Attitude Inventory than the control subgroup, two of them significantly.

SUMMARY TABLE 4

Comparison of Experimental Groups with Control Group (A) on Tests Not Primarily Dependent on Word Recognition (Total Treatment Groups)

Experimental Group	Stanford Paragraph Meaning	Stanford Vocabulary (oral)	Stanford Spelling	San Diego Attitude
B		x	x	x
C	x	x		x
D	x			x

In the total treatment groups no differences were significant though the trend is slightly in favor of the experimental groups. No unusual emphasis was put on the skills measured by the Stanford Paragraph Comprehension, Vocabulary, and Spelling subtests in any of the groups, experimental or control. It appears that there was some effect upon attitude in the total groups of which the experimental subgroups were a part.

In Summary Tables 5-8, an x indicates that the mean of one experimental group (as indicated in the lefthand column) exceeds the mean of the other experimental group. If the difference is significant, a 5% or 1% indicates the degree of significance.

SUMMARY TABLE 5

Comparison of Experimental Groups with Each Other on Tests Mainly Dependent on Word Recognition (Low Subgroups)

	Stanford Word Reading	Stanford Word Study Skills	Gilmore Accuracy	Gilmore Rate	Fry	Gates	Karlsen
Mean of B exceeds Mean of C				5%			
Mean of B exceeds Mean of D			x	x			
Mean of C exceeds Mean of B	x	x	x		5%	x	x
Mean of C exceeds Mean of D			x				
Mean of D exceeds Mean of B	x	x			5%	x	x
Mean of D exceeds Mean of C	x	x		x	x	x	x

The important fact in Summary Table 5 is that in the seven comparisons made between group B and group C the mean of group C exceeds the mean of group B six times. In both groups C and D the Houghton Mifflin readiness materials and the trade books were used. Group B had additional teacher time devoted to the low subgroups but used the usual basal materials. The trend of the scores indicates that the materials were more important to the pupils' performance than additional teacher time. Group D which had both additional teacher time and the different materials was superior to group C in six of the seven measures, indicating that a combination of the materials used and the additional teacher time may have been more effective than the materials alone.

SUMMARY TABLE 6

Comparison of Experimental Groups with Each Other on Tests Mainly Dependent on Word Recognition (Total Treatment Groups)

	Stanford Word Reading	Stanford Word Study Skills	Gilmore Accuracy	Gilmore Rate	Fry	Gates	Karlsen
Mean of B exceeds Mean of C							
Mean of B exceeds Mean of D							
Mean of C exceeds Mean of B	x	x	5%	x	1%	5%	5%
Mean of C exceeds Mean of D			x			x	
Mean of D exceeds Mean of B	x	x	x	x	1%	5%	5%
Mean of D exceeds Mean of C	x	x		x	x		x

A comparison of Summary Table 6 with Summary Table 5 reveals a marked similarity. It would appear that the presence of the different materials in the C and D classrooms was affecting the total group even though only the low subgroups were being directly taught with these materials.

SUMMARY TABLE 7

Comparison of Experimental Groups with Each Other on Tests Not Mainly Dependent on Word Recognition (Low Subgroups)

	Stanford Paragraph Meaning	Stanford Vocabulary (oral)	Stanford Spelling	San Diego Attitude Inventory
Mean of B exceeds Mean of C	x		1%	x
Mean of B exceeds Mean of D		1%	x	
Mean of C exceeds Mean of B		x		
Mean of C exceeds Mean of D		1%		
Mean of D exceeds Mean of B	x			x
Mean of D exceeds Mean of C	x		x	x

No special emphasis was placed on the skills measured by the Stanford Paragraph Meaning, Vocabulary, and Spelling subtests. No particular trend is observable in these results.

SUMMARY TABLE 8

Comparison of Experimental Groups with Each Other on Tests Not Mainly Dependent on Word Recognition (Total Treatment Groups)

	Stanford Paragraph Meaning	Stanford Vocabulary (oral)	Stanford Spelling	San Diego Attitude Inventory
Mean of B exceeds Mean of C			x	
Mean of B exceeds Mean of D		1%	5%	
Mean of C exceeds Mean of B	x	x		x
Mean of C exceeds Mean of D		1%	x	x
Mean of D exceeds Mean of B	x			x
Mean of D exceeds Mean of C	x			

No trend is observable in Summary Table 8.

The trends revealed in Summary Tables 1-8 may be interpreted tentatively to mean the following:

1. The use of the Houghton Mifflin readiness materials plus trade books in place of regular basal readers tends to improve the performance of children in grade one who are likely to have difficulty in learning to read, especially on measures which are heavily dependent on word recognition skills.

2. The use of remedial teachers to support the regular classroom teacher's work with children in grade one who are likely to have difficulty in learning to read also tends to

improve the performance of these children, especially on measures which are heavily dependent on word recognition skills.

3. The use of the Houghton Mifflin readiness materials plus trade books in place of regular basal readers also tends to improve the performance of all children in the classroom on measures which are heavily dependent on word recognition even though only the children likely to have difficulty are using these materials. There is less evidence that the use of remedial teacher time with the children likely to have difficulty has any effect on the room as a whole.

4. In both the low subgroups (children likely to have difficulty learning to read) and the total treatment groups the scores indicate that the use of the Houghton Mifflin materials plus trade books in place of basal readers is more effective than remedial teacher time spent with the low subgroups. However, the combination of the special materials with remedial teacher time appears to be more effective than either by itself.

While the total battery of predictive tests used at the beginning of grade one was effective in identifying the children likely to have difficulty in learning to read, Pearson Product Moment correlations between these tests and selected posttests are too low to warrant drawing the conclusion that any one or two of these tests are reliable predictors. In this respect, this study tends to support earlier prediction studies.

The study suggests one particularly interesting question for further research:

Does a prolonged and intensive readiness program concentrated upon basic skills with letter sounds actually make the controlled vocabulary usually associated with basal readers unnecessary (as the use of the trade books in this study suggests)?

APPENDIX A

TRADE BOOKS

Name	Author	Publisher
Are You My Mother?	P. D. Eastman	Random House
Come and Have Fun	Clement Hurd	Harper and Row
Funny Baby, The	Margaret Hillert	Follett
Gertie the Duck	Georgiady and Romano	Follett
Golden Egg Book, The	M. W. Brown	Golden Press
Have You Seen My Brother?	Elizabeth Guilfoile	Follett
Hill That Grew, The	Esther Meeks	Follett
Home for a Bunny	M. W. Brown	Golden Press
Hop on Pop	Dr. Seuss	Random House
Just Follow Me	Phoebe Ericson	Follett
Little Quack	Ruth Woods	Follett
My Own Little House	Merriman Kaune	Follett
Nobody Listens to Andrew	Elizabeth Guilfoile	Follett
Sad Mrs. Sam Sack	Brothers and Botel	Follett
Ten Apples Up on Top	Theodore LeSieg	Random House
Three Bears, The	Margaret Hillert	Follett
Three Goats, The	Margaret Hillert	Follett
Three Pigs, The	Margaret Hillert	Follett
Tiny's Big Umbrella	M abel LaRue	Houghton Mifflin
Tiny Toosey's Birthday	Mabel LaRue	Houghton Mifflin
Too Many Dogs	Ramona Dupre	Follett
Who Will Be My Friend?	Syd Hoff	Harper
Who Will Milk My Cow?	Janet Jackscn	Follett

Sample of teaching suggestions prepared for trade books

NOBODY LISTENS TO ANDREW

INTRODUCE BOOK

Where is the boy? Where is he pointing? Why do you think he is pointing upstairs? This little boy is called Andrew. He is a little boy who talks so much that people just don't listen to him. Let's read our title on the cover of the book. Tell us again why nobody listens to Andrew.

Turn to the title page in the book. Look at Andrew. What is he doing? Is anyone listening to him? What is everyone doing as Andrew talks? Let's read the title again. What else does the title page tell us? (Be sure children realize that the author, illustrator, publisher, and title of book are all given here.)

THREE RULES FOR EACH PAGE

1. Set the scene by studying each picture.
2. Have children read silently first with a definite purpose.
3. Follow with oral reading to show thought and understanding.

Pages 4 and 5

Andrew	upstairs	down
saw	he	very
something	ran	fast
Mother	listen(s)	said

Always introduce new words before beginning reading of a page. Some of the words listed may be old words depending on what trade books children have previously read. Flash card drill on known words before introducing the new words will prove beneficial at this time. Introduce all new words following the Houghton Mifflin method (see last page of this manual).

Suggested questions are given. Delete or add as you see the need.

Look at the picture that goes across both pages. Where is Andrew? How does he look? What is he doing? How old do you think he is? Do you think he goes to school as you do? Let's read to see why he is so excited.

Page 6

I	Mrs.	catch	before
wait	Cleaner	the	dark
pay(ing)	she	bus	

Andrew is still talking, isn't he? What is Mother doing? What do you think Mrs. Cleaner has been doing for Mother? Where does she have to go after she finishes her work?

Read this page to yourself and find out why Mother told Andrew to wait.

Read the sentence out loud to us that tells what Mother had to do. Read the sentence that tells what Mrs. Cleaner had to do. What word did the author use in that sentence to take the place of "Mrs. Cleaner"?

Page 7

Daddy	cut	grass

Mother wouldn't listen to Andrew so he has gone to tell someone else. To whom is he talking now? Is he still inside the house? What is Daddy doing? Do you think Daddy will listen to Andrew?

Read the first three lines and see what Andrew told Daddy. What word tells you what he saw? (something) What word tells you where he saw something? (upstairs)

Who is going to talk now? Finish the page and find out what Daddy told Andrew to do. Read the sentence that tells us what Daddy said. Now read the sentence that tells why Daddy couldn't stop and listen to Andrew.

Page 8

Ruthy	in	put	skate(s)
it	my	on	want
was	bed	roller	to

Who might the little girl be? What is she doing? Do you think she will stop and listen to Andrew? Why is he pointing toward the house?

Read the first four lines and see if you know exactly where Andrew saw something. Can you read one sentence that tells exactly where it is?

Finish the page and find out two things that Ruthy wants to do. Read one thing she wants to do. Now read another thing. Who else has something to do before it gets dark? (Mrs. Cleaner and Daddy)

Page 9

Bobby	sun	porch

Who is that other boy in the picture? Could it be an older brother? What do you think his name is? What is he doing? Does Andrew still seem excited about what he saw upstairs on his bed?

We know Andrew saw something on his bed. Now read this page to yourself and find out where Andrew's bed was.

Read out loud four words that tell where his bed was. Now read all that Andrew said to his brother Bobby.

Page 10

don't	me	bat	ball
bother	find	and	play(ing)

What is Bobby doing? Why would anyone throw everything out of a closet like that? What is Andrew doing?

Read this page to yourself and find out what Bobby wanted to do.

Read the sentence that tells what he told Andrew not to do. Now read the sentence that tells what he is looking for. Read the sentence that tells what Bobby wants to do before dark. Can you read all that Bobby said and make it sound as if he doesn't want to be bothered by Andrew?

Page 11

Mr.	black	dog
Neighbor	never	for
take	mind	walk

Where has Andrew gone now? Who is that man? What is he doing?

Read the first two lines and see who this man is. What did Andrew call him?

Read the next three lines and see if Andrew tells anything new about what he saw upstairs on his bed. Read out loud the line that tells us something new.

Finish the page and see what Mr. Neighbor must do before dark. Do you think he believes Andrew? How do you think he said, "Never mind, Andrew." Read what Mr. Neighbor said and make it sound as if he didn't believe Andrew.

Page 12

loud	there	is	bear(s)

Look at Andrew. What is he doing? What tells you he is talking very loud? What tells you he is stamping his feet?

Read this page to yourself and find out what is upstairs on Andrew's bed. Why is the last sentence on this page printed in big capital letters? Let's read it and make that part very loud when we come to it. Do you think everyone will listen to Andrew now?

Page 13

stopped	police	fire
call(ed)	cutting	department

Are the other people excited now? How do you know?

Read the first two lines and see what Mother stopped doing and whom she thinks they could call.

Read the last two lines and find out what Daddy stopped doing and whom he thinks they should call.

Page 14

catcher	skating	zoo

What do you see in the picture that tells you Bobby and Ruthy were excited when they heard what was on Andrew's bed? Notice the lines that tell you the ball and Ruthy were both going very fast.

Read the first two sentences and find out what Bobby thinks they should do.

Read the last two lines and see what Ruth thinks they should do. Do you think they need to call all those places and people?

Page 15

taking	his

What is Mr. Neighbor doing? Let's see if we can think of some of the places and people he might be calling.

Read this page and see how many of those we named Mr. Neighbor really called on the phone. Read us a sentence that tells us one place he called. Another? Another? Do you think they will all come? Was Mr. Neighbor excited?

Pages 16 and 17

zoom	came

Look at the picture that goes across both pages. Are these men policemen? (Make sure children realize that some policement walk, others ride horses, motorcycles or in police cruisers.) Are they all going fast?

Read these pages and find one word that tells how fast they are going.

Pages 18 and 19

zing

How many men are riding on the firetruck? What do the letters FB stand for? What is that word on the other red car? Who is riding in it? Does that dog belong with the firemen? Why? Are they going as fast as the policemen?

Read silently and find a word that tells how fast the firemen are going.

Pages 20 and 21

whoosh	man	they
swish	from	all

Look at the picture that goes across both pages. Where are all those people? Can you find the policemen? Firemen? Dog catcher? Zoo men? Where are they going? Will there be anything upstairs on Andrew's bed?

Read the first line and find out what noise the dog catcher made as he came running into the house.

Read the second sentence and find out what noise the man from the zoo made.

Read Page 21 and see where they all went. What word tells us?

Pages 22 and 23

look(ing)	but

Look across both pages. What was really on Andrew's bed? Was Andrew telling the truth when he said there was something black on his bed? Should his family have listened to him? Do you think they will listen to him after this?

Read the first two sentences and find out what Mother said when she saw the bear.

Read the next line and see what Daddy said.

Read the next two sentences and find out what Bobby said about the bear.

Read the rest of the page and see what Ruthy said.

Dramatize this page with children taking parts of characters.

Page 24

caught	fireman	up	window

Does this seem like a nice friendly bear? Does he seem surprised?

Read the first sentence and see who caught the bear.

Finish reading the page and find out how that bear got on Andrew's bed. (Turn back to previous page so children can see how it could climb in the window after climbing the tree.)

Page 25

dry	are	water	this
woods	thirsty	will	

Who is that man? What has he put over the bear's head? Why?

Read the first two lines and see what the man from the zoo said about the woods.

Read the next two sentences and find out what he thought the bears were looking for.

Read the last sentence and see where he is going to take the bear. Is that a good place for wild bears?

Page 26

next	time	we

How do you think the family feels about Andrew now? How do you think Andrew feels about it?

Read this page to yourself and find out if the family has changed its opinion of Andrew. Are they going to listen to him after this?

This book may be reread orally more than once for plain enjoyment. It is easy and effective to dramatize. It can be the starting point for a unit on community helpers.

SKILLS USED IN READING THIS BOOK

Beginning consonant sounds combined with context clues to unlock new words.

Beginning digraphs combined with context clues to unlock new words:

th - the, there, they, thirsty, this
wh - whoosh

Beginning consonant blends and context clues:
cl - cleaner; gr - grass; sk - skate, skating; pl- playing; bl - black;
st - stopped; sw - swish; fr - from; cl - climbed

Compound words: upstairs, fireman

Endings on root words: listen(s), pay(ing), skate(s), play(ing), bear(s), call(ed), look(ing)

Antonyms: upstairs, downstairs; down, up; fast, slow; before, after; play, work; loud, soft.

Contraction: don't

EXERCISES THAT CAN BE USED THROUGHOUT THE READING OF THE TRADE BOOKS

Unlocking strange words
Print pig on the board. Then say, "There were four animals in the barn - a horse, a cow, a lamb, and a _____."
Ask: How do you know it isn't pony? (Pony doesn't end with the sound the letter g stands for.) How do you know it isn't paste? (Paste doesn't make sense.) How do you know it isn't dog? (Dog doesn't begin with the sound made by the letter p.)

(Notice that the final sound of a word is now being used as well as the beginning sound.)

Riddle Game
Print p on the board. Let's play a special kind of guessing game. I will tell you something about a word I am thinking of that begins with the sound p stands for. You see if you can tell me what word I am thinking of.

I am thinking of an animal that children like to ride. (pony)
I am thinking of something you sprinkle on food. (pepper)
This word is the name of something you write with. (pencil)
You will use this when you want to stick things together. (paste)
These are very good to eat for breakfast. (pancakes)

Play this game each day. Cover a different consonant, digraph, or blend.

Picture Cards
Place picture cards for four or five different consonants in the pocket chart. Then say: Find the picture that begins with the sound made by the letter b.

Continue in this fashion until all the pictures are removed from the chart. You can also say: Find the picture that begins like look and little. Find the picture that begins with the same sound as door and doctor.

Rhyming Game

Let's play a rhyming game. I will say all except the last word of some jingles. Every time I stop, I will ask some one to finish the jingle. The word that you choose must rhyme with a word that I print on the board. Print can on the board. Say your rhymes in pairs of sentences or lines. After you say the first line of each jingle, erase the beginning consonant of the word on the board and substitute the consonant that will make the sound in the word that the children will use to complete the jingle.

Take the c away from can;
Put in m and you have man.

Substitution

Print can on the board. While children watch, erase the n and put in t. Now the children should use the sound made by the letter t to make the word cat. Make whatever words are sure to be in the speaking vocabulary of the children.

Word Cards

Place word cards on the chalk rail. Then say, "Bring me the word that begins like new and no. Bring me the word that begins like jump and juice." Continue until all words are removed from the chalk rail. Use whatever single consonants, digraphs, or blends need the extra drill. Carry this type of work over into the final sound. Say: Find the word that ends like dog. Find a word that ends like wish.

APPENDIX B

PHONETICALLY REGULAR WORDS ORAL READING TEST

1966 Version

Edward Fry, Rutgers University
New Brunswick, N. J.

Name_____ Date_____

School_____ Room _____ Code Number_____

Examiner_____ Number of words read correctly _____

1. nap	16. stalk	31. yoke
2. pen	17. haul	32. glory
3. hid	18. jaw	33. shy
4. job	19. soil	34. quaff
5. rug	20. joy	35. taught
6. shade	21. frown	36. bundle
7. drive	22. trout	37. nix
8. joke	23. term	38. civic
9. mule	24. curl	39. Philip
10. plain	25. birch	40. preach
11. hay	26. rare	41. cracked
12. keen	27. star	42. swish
13. least	28. porch	43. frankfurter
14. loan	29. smooth	44. twelfth
15. slow	30. shook	45. drowse

Directions to Examiner: Have pupil read words from one copy while you mark another copy. Do not give pupil a second chance, but accept immediate self-correction. Let every pupil try the whole first column. If he gets two words correct from word number six on, let him try the whole second column. If he gets three words correct, let him try the whole third column. Mark correct words C and incorrect words X.

GATES WORD PRONUNCIATION TEST

EXAMINER'S COPY

DIRECTIONS: Have the child read the words out loud. Tell him you would like him to read some words for you. If he fails the first time, ask him to try the words again. Continue until ten consecutive words have been missed. As the words become difficult, special care should be taken to encourage the child. The score is one point for each word correctly pronounced on the first trial, one-half point for each word correctly pronounced on the second trial. (Note: 9 1/2 correct would be scored as 10.)

co

---

1. so
2. we
3. as
4. go
5. the
6. not
7. how
8. may
9. king
10. here
11. grow
12. late
13. every
14. about
15. paper
16. blind
17. window
18. family
19. perhaps
20. plaster
21. passenger
22. wander
23. interest
24. chocolate
25. dispute
26. portion
27. conductor
28. brightness
29. intelligent
30. construct
31. position
32. profitable
33. irregular
34. schoolmaster
35. lamentation
36. community
37. satisfactory
38. illustrious
39. superstition
40. affectionate

Child's name _____ Test date _____

Examiner _____ Birth date _____

Age _____

KARLSEN PHONEMIC WORD TEST

Examiner's Copy

Directions:
1. Hand the PUPIL'S COPY to the pupil.
2. Say: "Read these words out loud."
3. Note the pupil's errors on this sheet.
4. Do not give the pupil a second chance, but accept immediate self-correction.
5. Continue until the child misses 5 consecutive words.
6. The score is the number of words pronounced correctly.

---

1. fit
2. tap
3. rod
4. get
5. would
6. mother
7. down
8. age
9. think
10. long
11. kind
12. yard
13. foot
14. gold
15. freeze
16. chair
17. mouth
18. carry
19. hope
20. beat
21. loaf
22. cowboy
23. furniture
24. page
25. push
26. huge
27. snowball
28. thirteen
29. scare
30. sunshine
31. gymnasium
32. join
33. usual
34. teaspoon
35. zone
36. monument
37. senior
38. flute
39. behave
40. faucet

Child's name: _____ Test date _____

Examiner: _____ Birth date _____

Age _____

INFORMAL TEST OF LETTER WRITING

Directions for teachers:

Each letter should be made within the box. Any recognizable letter, capital or small, will be scored correct.

In giving the test, say: Put your finger in the first box. (Show them which is first.) Now make the letter H in that box. Make the letter H. (Say each letter twice.)

Allow a reasonable time. Then say: Move your finger to the next box. Make K in this box. Make K.

Proceed in the same way for other boxes. Watch to see that children make the letters in the right boxes. If you think children will get mixed up, make a copy of this page on the board and point to the box you want the children to use.

Children will use pencils. Please see that the name of each child is on the back of his paper. This could be done before the test period.

1. Make the letter H.
2. Make the letter K.
3. Make the letter B.
4. Make the letter O.
5. Make the letter Y.
6. Make the letter S.
7. Make the letter N.
8. Make the letter M.
9. Make the letter A.
10. Make the letter W.
11. Make the letter R.
12. Make the letter F.

ANSWER SHEET

1.	2.	3.	4.
5.	6.	7.	8.
9.	10.	11.	12.