A DULUTH TENTH GRADE READING PROGRAM--A FOLLOW-UP STUDY

A Paper
Presented to
the Faculty of the Graduate School
University of Minnesota

Problems in
Curriculum Construction
Ed. C. I. 271
Under the Direction of
Harry C. Johnson

A Requirement for the Degree

Master of Arts (Plan B)

---

by
John Thomas Ronan
University of Minnesota
Duluth, Minnesota
August, 1967

TABLE OF CONTENTS

LIST OF TABLES

CHAPTER I

INTRODUCTION

The formal reading program as an integral and clearly identifiable entity in the secondary school is relatively new. At the senior high school level, such courses, though no longer rare, are even more recent and far from prevalent.

A six-week formal reading course for tenth-graders has existed at East High School, Duluth, Minnesota, for some time.

During the 1962-1963 school year this program was taught and studied by Edward G. Summers, now Dr. Summers, a staff member of the University of Indiana.

Summers did two studies, which have been published, of samples of the tenth-graders he taught.[1] In addition, the test scores of all students were preserved. This preservation of tenth grade scores makes the present study possible.

The present study attempts to answer two questions--What is the effect of the reading program on these students'

---

[1]Edward G. Summers, "Evaluation of Reading Gains in a Secondary School Reading Laboratory," The Reading Teacher 17:255-259, January, 1964; Edward G. Summers, East High Reading Center Report. Duluth, Minnesota; Duluth Public Schools, November, 1962.

reading skills two years after completing the reading course described in the published reports? Do the published reports of progress for samples of these students accurately portray the overall successes and failures of the reading program?

This study, like the original studies, cannot make absolutely definitive statements about the success or failure of the students in the program compared to students in another reading program or no program. There was no tenth-grade group at East High School which did not have reading instruction. Therefore, there is no control group for this study, just as there was no control group for the original studies.

This paper, then, is intended to provide further clues as to the effectiveness of the type of program reported upon by Summers. The problem of evaluation is approached by an analysis of pre and post program scores on a standardized test and results of a follow-up test administered two years later. In addition, samples of these students' changes in test scores are correlated to determine the degree of relationship existing between changes in reading skill occurring during the period of reading training and reading skill changes occurring throughout the first two-and-one-half years of senior high school.

Before describing the 1962-1963 tenth-grade reading program at East High School, a review of the published studies of similar programs would be appropriate.

CHAPTER II

REVIEW OF RELATED RESEARCH

A number of reports on high school reading programs have been published since the end of World War II. The Journal of Reading has carried an annual bibliography of published studies for the past several years. More recently, the Quarterly Review of Reading Research, published by the International Reading Association, has extensively publicized the research in reading at all levels. Previous to the 1960's, the Review of Educational Research, February issue, was the most up-to-date source of abstracts on reading research. The Review of Educational Research abstracts are still being published annually in the field of reading. The concise, yet richly descriptive, abstracts found in this source make it one of the finest bibliographical sources available.

Obviously, the scope of the review in this paper must be limited. The studies included in this review were chosen because they bear a similarity in some manner to the province of the present study. Reports are included on the basis of a relationship to the program, type of student, or grade level involved in the present study. In a few cases, reports are included primarily for purposes of comparison

and contrast.

The Summers[1] studies are not included in this chapter. They will be dealt with in greater detail in chapter three. The program evaluated in the present study will be much more closely scrutinized than will be the case with the studies reviewed in this chapter.

Part I--Similar Programs

Perry[2] reports a program of nine-weeks duration in Highland Park, Illinois, in which students might be considered fairly similar to those in this study. Priority was given to students of average, above average, and very superior reading ability.[3]

Highland Park had a reading center where students came to the class from a study period every day for nine weeks. Classes were limited to ten to fifteen students. Although Perry states that enrollment was voluntary, his report indicates that at least some psychological pressure was

---

[1]Edward G. Summers, "Evaluation of Reading Gains in a Secondary School Reading Laboratory," The Reading Teacher 17:255-259, January, 1964; and East High Reading Center Report. Duluth Public Schools, November, 1962.

[2]Harold J. Perry, "The Developmental Reading Program in the Highland Park High School, Highland Park, Illinois," pp. 122-127 in Elizabeth A. Simpson, Helping High School Students Read Better. (Chicago: Science Research Associates, 1954).

[3]Ibid., p. 123.

exerted by counselors where it was felt students who should have enrolled had not done so.[4]

Prominently mentioned in this report was the use of accelerators and the tachistoscope. Materials used included Gray's Basic Reading Skills for High School Use, Reading for Meaning, and SRA Better Reading Books with Progress Folders.

During the spring of the first year of the program (1951-1952), standardized tests were administered to see if students who left the reading center earlier in the year retained the gains made. Perry states, "The test results indicated that these students had retained nearly all of their reading gains."[5]

Several comments are in order. The smallness of the reading classes is unusual for a developmental reading program. Although far larger than the ideal remedial class, the number of students handled at one time is much closer to that actually found in remedial groups than to the usual size for developmental reading classes.

The priority given to average and above average readers is found fairly commonly in reports of programs of the early 1950's. There seems to be a reversal of the fairly common

---

[4]Ibid., p. 124.

[5]Ibid., p. 126.

view of that time that the below-average were beyond hope when they had attained high school grade status. Although high schools often have developmental programs which do not seem very appropriate for slower readers, more recent reports rarely indicate that poorer readers are excluded from the reading program.

As is often and regrettably the case, the results of the reading program at Highland Park are vaguely reported; and the usage made of the materials is not fully explained. Drawing comparisons about the relative effectiveness of various reading programs is, therefore, usually impossible.

Very similar to the Highland Park program was one tried at Oak Park, Illinois.[6] Here the class was also limited to fifteen students. Students came from study halls and enrollment was voluntary with no grade or credit given. Students attended for a minimum of twenty sessions.

The approach used appears to have placed great emphasis upon individualization and to have been very similar to a remedial approach. Based upon diagnostic tests, students were assigned areas of work. Some might work only on speed (accelerators), while others worked only on vocabulary, for

---

[6]Kermit Dehl, "Reading Improvement Services in the Oak Park and River Forest High School, Oak Park, Illinois," pp. 118-121 in Simpson op. cit.

example.[7]

Dehl reports that these students made good progress. A gain of 50 to 400 words per minute (WPM) at 80 to 100% comprehension in factual-type reading, along with even greater gains in novel reading, is claimed for students who had worked on rate improvement. Those who worked on vocabulary or general comprehension usually gained 10 to 25 points in percentile rank on standardized test norms.[8]

Dehl does not name the tests used to arrive at these indications of improvement, nor does he indicate the time interval which might elapse between a student's first and last sessions in the reading laboratory. Mental and reading ability characteristics of the students who took the reading class are not mentioned.

The Evanston Township High School (Evanston, Illinois) reading program was reported at two different times in two different sources.[9] The earlier report will be dealt with in detail. Variations from the earlier report found in the

---

[7] Ibid., p. 119.

[8] Ibid., p. 121.

[9] Phyllis Bland, "The Development of a Reading Program in the Evanston Township High School, Evanston, Illinois," pp. 107-115 in Simpson op. cit.; Phyllis Bland, "A High School Developmental Reading Program," The Reading Teacher, 8:146-52, February, 1955.

later publication will then be noted.

The Evanston school made what appears to have been a successful attempt to provide a comprehensive reading program for grades nine through twelve. Grade nine was used for emphasis in remedial reading for those who needed it. This was accomplished by designating certain English sections for that purpose. The ninth grade program will not be discussed here as it does not relate to this paper. In grades ten through twelve a two-pronged approach was used.

Reading skill classes were elective and open to all students. These students attended reading class every other day all year and received one-half credit. How to Become a Better Reader by Witty was the basic text for this class. SRA Better Reading Books were used as were a flashmeter and reading accelerators.

The accelerated reading class was a six-week course open to average or above-average readers who wanted to get better. A wide range of materials on various aspects of reading was prominent in this class. Mechanical devices were not mentioned for the accelerated group.

The effectiveness of the reading program was measured by pre and post tests. The Diagnostic Reading Tests, Survey Section was the instrument used. The report of the results for the skill and accelerated groups was combined. In rate

there was an increase from 283 to 370 WPM, in comprehension of story-type material there was an increase from 42nd to 80th percentile rank, in comprehension of main ideas and details there was an increase from 33rd to 53rd percentile rank.[10]

In the second Bland report most factors were the same. However, rather than meeting every other day all year, skill classes reportedly met on a semester basis.[11] It seems fair to assume that this descrepancy was caused by a desire for brevity in the later report and that the program was operated on an every other day basis. This assumption is based on the fact that results for the program were exactly identical in the two reports. They must, therefore, have been separate reports of an identical study.

Another program which utilized time taken from study halls has been reported by Tormey and Patterson.[12] Students in this program went to the reading center two days per week. Some attended for one semester, others for the entire year. Classes were limited to twenty students per session

---

[10]Bland, "The Development etc. ...," op. cit. p. 115.

[11]Bland, "A High School etc. ...," op. cit. p. 150.

[12]Mary K. Tormey and Walter G. Patterson, "Developmental Reading and Student Evaluation," Journal of Developmental Reading, 2:30-43, Winter, 1959.

and the reading center was staffed by two teachers. The physical set-up was rather unique in that two adjoining rooms were utilized with one teacher in each room. Apparently, one room was used exclusively for mechanical devices and the other room was used for textual reading aids.

Mechanical devices used included tachistoscope, Controlled Reader, and ten accelerators. Materials mentioned in the article included: SRA Better Reading Books, Better Work Habits by Salisbury, How to Study by Preston and Sotel, How to Improve Your Reading and How to Become a Better Reader by Witty, and an SRA Reading Laboratory.

The program was evaluated on the basis of percentile rank change on alternate forms of the Iowa Test of Silent Reading. Of the 410 students in grades nine through twelve who were in the reading center for a semester or a year, 26 raised their percentile rank by more than 50; 197 showed a gain between 21 and 50 percentile points; 160 showed a gain of 1 to 20 percentile points; 16 showed neither gain or loss of percentile points; 9 showed a loss of 1 to 10 percentile points; 3 showed a loss of 11 to 20 percentile points; and 1 showed a loss of between 21 and 30 percentile points.[13]

---

[13] Ibid., p. 40.

It should be pointed out that the first test in September and the second test in May would find students in the same grade placement. Some percentile rank improvement would be almost inevitable for a group over this period of time. This is not meant to imply that the gains do not seem impressive. As appears to be the case with most studies of reading programs, the validity of these results is clouded by the lack of a control group. This shortcoming is not present in the following study.

A high school in Atlanta, Georgia, matched two groups of 40 seniors each.[14] For one semester the experimental group emphasized reading in English class. Reading accelerators were used to work on rate improvement. Materials included Robinson's Effective Study, Witty's How to Become a Better Reader, and SRA Better Reading Book 3. Special reading materials and accelerators were not available to the control group.

On the basis of alternate forms of the Iowa Silent Reading Test administered at the beginning and end of the semester, both groups showed significant gains from pre to post testing. The experimental group gained significantly

---

[14]Cecil M. Thornton, "Two High School Reading Improvement Programs," Journal of Developmental Reading, 3:115-22, Winter, 1960.

more than the control group on rate. Although the difference was not significant, the control group gained more than the experimental group on vocabulary and comprehension.

Thornton concludes that the teacher of the experimental group relied too heavily upon mechanical devices for speed increase when the students were more in need of emphasis upon other skills.[15]

Melton reports a follow-up study which was very similar to that found in this paper.[16] The description of the six-week, tenth-grade program is quoted in its entirety.

> The program consists of concentrated attempts to improve reading speed and comprehension. Specified textbooks, SRA workbooks 1, 2, and 3, and Iowa Silent Reading films and tests are used in the reading experience.[17]

Alternate forms of the Iowa Silent Reading Tests were used for testing before and after reading training and in the follow-up test. The follow-up test was administered during the fifth six-weeks of the eleventh grade. The reading training took place at any time during the tenth grade. So, the follow-up could probably best be thought of as having taken place about one year after the close of

---

[15] Ibid., p. 121.

[16] Howard B. Melton, "Reading Retention of Eleventh-Grade Students," Journal of Developmental Reading, 2:26-33, Summer, 1959.

[17] Ibid., p. 26.

reading training. Results were reported only on the basis of percentile rank. The results are for a random sample of eleventh graders, not the entire group which took the reading course.

From pre to post test the median percentile rank rose from the 45th to the 75th percentile. On the follow-up test the median percentile rank for the sample was 68. From pre to follow-up test, 84% of the sample gained percentile rank. From post test to follow-up test, 42% of the sample gained in percentile rank.

Means, medians, and percentile rank can all be criticized as avenues of comparison in evaluating a program; and the preceding analysis certainly can be criticized. Nevertheless, the fact that 42% of the students measured a gain in percentile rank from post to follow-up test is outstanding. Practically all reading programs of short duration, whether remedial or developmental in nature, show more regression than that found in this study; and very few can claim that this high a proportion of students continue to improve at a rate beyond the normal growth curve.

A program very similar to that found at East High School, Duluth, is described by Kendall.[18] As described,

---

[18]Rose Kendall, "Give Them a Chance," Journal of Reading, 8:326-9, April, 1965.

this program placed great emphasis upon speed, with vocabulary the only other area of reading touched. Evaluation of the program was based upon the same test as that used in the present study. Kendall states that approximately 75% of all students showed definite improvement.[19] The progress of a particular class was described as follows:

> In one class of twenty-six students, all except one increased their reading rate at least two hundred words per minute and improved comprehension at least ten percent.[20]

The gains claimed in the above statement seem very impressive and they may be. However, the testing instrument, Diagnostic Reading Test, Survey Section, allows such statements to be made without being very meaningful. There are two sub-tests which, together, constitute the comprehension score. An increase of ten percent in comprehension does not necessarily mean that the improved comprehension was associated with the higher rate of reading. If gain on the other sub-test were greater, comprehension on the timed reading sub-test could have dropped from pre to post test and it would still be possible to say that overall comprehension had increased. For this reason, the results

---

[19] Ibid., p. 329.

[20] Ibid., p. 328.

reported in this article must be considered ambiguous.[21]

A very short report on a program which emphasized speed almost exclusively may be found in the April, 1965, edition of the Journal of Reading.[22] This was apparently a semester course at the senior high level in which reading films were shown for one-half the period and Shadowscope pacers were used during the other half of the period.

On alternate forms of the Triggs Diagnostic Tests, students showed a pre to post rate gain of 109 words per minute--from a mean of 251 to a mean of 360 WPM. On total score the mean rose from 73 to 76. Tremendous gains were registered from the beginning to the end of the course on speed in reading training films and speed on the pacers.

The speed increases outside of the reading test seem somewhat irrelevant. The testing situation certainly more closely approximates the everyday reading situation than does the environment of mechanical aids.

The gains shown by the test results are not startling considering the great emphasis put on speed. It would be more amazing if the gains were small under the reported

---

[21]See chapter five for a more complete discussion of the ambiguous implications of the term comprehension in relation to this test.

[22]Reata M. Engelhardt, "Speed Is Not a Naughty Word," pp. 330-31.

circumstances.

Berkey reports a ninth-grade program which utilized a period of laboratory training followed by continued reading emphasis in the English class.[23] The laboratory had a capacity of fifteen students. Every ninth grader had to attend for eight weeks from time ordinarily spent in English. After the laboratory experience, each student spent one or two days a week in supervised classroom reading practicing what he had learned in the reading laboratory.

On alternate forms of the Nelson Silent Reading Test these students had a mean grade score of 8.2 at the beginning of the year, 9.0 at the end of laboratory sessions, and 9.9 at the end of the school year.

It is regrettable that the entire program is not reported in more detail. The students made good progress in the reading center; but the continued growth after leaving the reading laboratory is far more interesting and, as previously noted, not commonly reported. Obviously, there is something very useful about the approach this program took; but, other than follow-up work, the article gives no clue as to the nature of that something which seemed to work so well. Neither the nature of the work done

---

[23]Sally Berkey, "Reading and Study Skills Program in a High School District," The Reading Teacher, 16:102-103 November, 1962.

in the reading laboratory nor the nature of the work done in the English classroom is disclosed.

East Bakersfield High School, California, developed a variable program for students of different grade and proficiency levels.[24] Only those aspects of the published report which relate to tenth graders are summarized here. The program featured a good deal of emphasis upon reading machines. The Controlled Reader and SRA Reading Accelerator were used extensively. In addition, however, a very wide range of multi-level reading materials were available to the students. Reading training covered a nine-week period.

On alternate forms of the Iowa Silent Reading Test, the mean growth for three groups (number of students not stated) was 1.53 years during the nine-week period. Therefore, the author reasons:

> . . . . . .For such a testing period progress in the reading improvement classes was approximately nine times the normal progress expected of typical students.[25]
> . . .

Two points in this evaluation deserve elaboration. It is not stated whether the 1.53 years mean growth was determined by averaging raw scores and then determining growth

---

[24]Ruth Freed, "Developing a High School Reading Laboratory," Journal of Secondary Education, 40:303-309, November, 1965.

[25]Ibid., p. 308.

in terms of grade placement or by averaging individual students' grade placement scores. If the former procedure was followed, the growth claim has more validity. Grade placement scores are not averagable, as such, because the units of such a scale do not indicate equal raw score values.

The statement that students gained nine times expectancy can be misleading in two ways. Most students in a concentrated reading program gain impressively over that period of time. Without a later test, it is impossible to know whether the gain will be retained or whether normal growth will be resumed after the period of reading training. There was acceleration of learning taking place in this program, but whether this was only pre-empting learning which would have taken place anyway over a longer period of time is not known.

In 1951-1952, the Iowa Reading Films were used in all the ninth grade classes of two high schools in Rochester, New York.[26] 473 students were included. The fourteen films were shown at the beginning of the class period at the rate of one per week for one-half the groups and two per week for

---

[26]Robert F. Barry and Paul E. Smith, "An Experiment in Ninth-Grade Reading Improvement," Journal of Educational Psychology, 45:407-414, November, 1954.

the other half. Each film was followed by a ten question comprehension check and students kept a record of their scores. During the same class period, students read a Reader's Digest article and kept track of reading rate for the fourteen articles read.

Immediately before, immediately after, three and one-half months after, and approximately fifteen months after the end of reading emphasis, students were tested on alternate forms of the Nelson Silent Reading test. The results are interesting.

There was no significant difference between the two groupings. During the experiment, the mean gain for all students was fifteen months on Nelson test norms (8.8 to 10.39). During the year interval following the third test the mean gain was only four months (10.4 to 10.8).

In 1952-1953, eight different approaches were used over a seven-week period for all freshmen in the Rochester school system.

Three groups used no reading films, one used only reading films, one used the same plan as the previous year, two used a lesser number of films along with textual material, and one group was only tested.

The first test was administered immediately before reading training. The second test was administered three

and one-half months after the end of the seven-week period. The results provide food for thought.

Although none of the approaches could be considered comprehensive, all groups showed good gains and there was very little difference from one approach to another. Two groups had a mean gain of 1.1 years on Nelson test norms during this interval; four groups (including the group that was only tested) had a mean gain of 1.2 years; one group had a mean gain of 1.3 years, and one group had a mean gain of 1.4 years.

This study will receive further comment at the end of this chapter.

Blough describes a twenty session course in which six sessions were devoted to explaining procedures, testing, or interpreting tests.[27] In the remaining fourteen sessions one Iowa Silent Reading Film was shown each day. During the remainder of the period students read essays and used pacers on alternate days. The grade level of these students is not stated. 324 students took the reading course during the first semester and 283 served as a control group. Students were tested in September, January, and May of the same

---

[27]Richard Blough, "The Developmental Reading Laboratory--A Training and Proving Ground for Readers," Bulletin of the National Association of Secondary School Principals, 40:38-41, November, 1956.

school year. Triggs Diagnostic Reading Test, Forms A, B, and C were administered.

In September the control group and the experimental group were very even in both rate and comprehension, the only results reported. The comprehension means were only .1 of a point apart and the mean rates were within 10 WPM. In January the spread in comprehension was 2.1 points in favor of the experimental group. In May the spread was over 5 points in favor of the experimental group. (The control group mean dropped 1.9 points between January and May.) The September to January gain in rate for the experimental group was approximately 37 WPM while the control group did not change. On the May test the experimental group showed an additional gain of approximately 20 WPM. At this point the control group showed a total gain of 5 WPM.

Without any statistical analysis of significance, the scores do not have much meaning. It is almost impossible to judge the effectiveness of the approach described.

Bish describes another course in which speed emphasis was strongly dominant.[28] 24 students in two classes were included in this study. The class was of nine-weeks

---

[28]Charles E. Bish, "An Experiment in Reading Improvement," Bulletin of the National Association of Secondary School Principals, 36:89-96, January, 1952.

duration, one forty-five minute session per day. Tachistoscope and accelerators were the mainstays of this program. There were some vocabulary exercises and outside reading was urged. All students in the program were of average or above-average ability.

The students were tested before, immediately after, and nine weeks after the end of reading training. Alternate forms of the Diagnostic Reading Tests, Survey Section, were administered. From pre to post testing the students measured significant gains in speed, vocabulary, and total reading. They showed a non-significant loss in story-type comprehension at the higher rate and a non-significant loss in total comprehension. On the follow-up test nine weeks later, students showed a significant increase in rate, a non-significant loss in comprehension of story-type reading, and a non-significant loss in vocabulary, a non-significant gain in comprehension, and a probably not significant loss in total reading when compared with pre test scores.

The results seem to show that this program succeeded in increasing speed while possibly lowering comprehension a little. Nothing else happened.

The previous study is especially important to the review for this paper, not because the program had great success, but because the same test instrument was used in

the present study. Even though the follow-up is much longer for the present study and the program studied here was more comprehensive, pre, post, and follow-up test scores exhibit similar tendencies for both programs. The appropriateness of the test instrument will be discussed further in the last chapter of this paper.[29]

An experiment was conducted in a large suburban high school with two classes of gifted tenth grade English students.[30] The control group was a class of 27 of these gifted students, and the experimental group consisted of 34 gifted students. The total description of the reading program is quoted below.

> The program for the experimental subjects consisted of a six-week intensive training on remedial reading techniques with emphasis on speed, consisting of lectures, reading materials, individual help, drill, tests, and the use of the tachistoscope and slides.[31]

On rate and composite scores of the Iowa Silent Reading Test taken before and after the program, the experimental group showed significant (beyond the .01 level of confidence) improvement from pre to post testing. The excess of the experimental group's improvement over the control

---

[29] See Limitations of the Study in chapter six.

[30] J. C. Gowan and R. W. Scheibel, "The Improvement of Reading in Gifted Children," Educational Administration and Supervision, 46:35-39, January, 1960.

[31] Ibid., p. 38.

group's improvement was at the .01 level of confidence.

The authors of this report conclude that the experiment illustrates how much gifted children can improve their reading rates and therefore reading adequacy by teaching them with remedial methods.[32] How they arrive at the conclusion that the program described was remedial in nature is a mystery.

There is nothing even remotely remedial about large groups, speed training, and lectures on reading methods. The program appears to have produced good results, but it should be considered a developmental reading program.

The final program reviewed in this section is really picked for contrast. It bears little similarity to the program taught at East High School. It seems wise to review a study where speed was not a factor in the program to see if the ambiguity so often found in the literature exists in a report of such a program.

This experiment involved a control class and an experimental class of senior English students.[33] The experiment was conducted over a twenty-seven week period. All factors were held constant except that the experimental

---

[32] Ibid., p. 38.

[33] Jeanne R. Jackson and H. Dizney, "Intensive Vocabulary Training," Journal of Developmental Reading, 6:221-9, Summer, 1963.

class devoted thirty-five class periods to intensive vocabulary training. Both classes covered exactly the same material, except for the vocabulary work, with the same teacher. The control group just took more time to cover the material in order to fill in the time the experimental group spent on vocabulary.

At the end of the experiment, after accounting for unavoidable differences in the groups, the control group was found to be insignificantly higher in speed and comprehension. The experimental group was higher (.05 level of confidence) in vocabulary.

The authors conclude that knowledge about the influence of vocabulary training on reading ability was not furthered by their experiment.

Obviously, not all the programs described in this section have been similar only in program approach. Some are for tenth graders only, but they have been classified as more similar in approach than in pupil personnel. The boundaries are somewhat artificial; many of the studies reported would fit under any of the three major divisions of this chapter.

---

[34] Ibid., p. 228.

Part II--Similar Students

The reviews in this section deal with students of superior academic ability. It is in this sense that they are considered similar to the students which are the subjects of the present study.

A program designated as pre-college is described and evaluated by Dobrin.[35] This program was voluntary and enrolled juniors and seniors for two days a week from their free hour. It was a sequential, four part program. Seniors who did not take the class as juniors, for example, would start in the same class as beginning juniors. Students could take the class for a semester or for the full year. Those falling below the 25th percentile on entering college freshmen norms received an additional day per week of remedial work with the instructor.

The vocabulary section of the program emphasized synonyms, antonyms, homonyms, figures of speech, word analogies and special terminology. Phonics was taught by rule to those who seemed to need phonetic instruction. Speed was increased through the use of the Controlled Reader and timed readings. Accelerators were available

---

[35]Ruth M. Dobrin, "The Massapequa Story: a Pre-College Reading Program," Journal of Developmental Reading 4:159-172, Spring, 1961.

for those who wished to come in an extra day during the week. Skimming, the survey study technique, and paragraph structure and pattern were systematically studied. (Presumably these areas of study were aimed at improving both speed and comprehension.)

From May of 1959 to January of 1960, a control group of students not in the program had a mean gain of 14.3 points on the College Board examinations. The groups of students who took the reading class during that period of time had a mean gain of 49.12 points on these tests.[36] Students in the program showed significant gains in vocabulary, rate, and comprehension.[37]

Several points stood out in this program. There was a strong emphasis upon a sequential as well as an individualized approach. The program was, or could be, continuous in nature. Students were not expected to change overnight and be good readers forever after. Probably most significant was the comprehensive nature of the instruction given in the program. It was not assumed that transfer of skill gain in one area of reading to another area of reading would occur; neither did the program assume that

---

[36] Ibid., p. 168.

[37] Ibid., p. 168.

any particular area of reading would be handled in a class other than reading.

A very carefully controlled, long-term study is reported by Glock and Millman.[38] Subjects were a group of high school juniors. The experimental group received two days per week of instruction in reading and study skills during the first semester of the eleventh grade. The control group received no special instruction. The program emphasized listening, note-taking, practice in rapid reading, and practice in areas of comprehension. Considering the breadth of the program, the results are somewhat surprising.

The experimental group did not show superiority in anything except reading speed (8% to 15% better than the control group) on follow-up evaluations in twelfth grade and college freshmen year. One of the measures of evaluation was grades received by students. It is interesting to note that there was no difference here either; the increased rate of reading did not have either a deleterious or enhancing influence upon achievement.

These two studies of superior students have somewhat

---

[38]Marvin Glock and Jason Millman, "Evaluation of a Study Skills Program for Above Average High School Pupils," Journal of Developmental Reading 7:283-9, Summer, 1964.

opposite results. It should be added that both programs were quite different than the program which is the subject of the present study and that both were longer in duration.

Part III--Tenth-Grade Programs

Many of the programs reported in previous sections have dealt mainly with tenth graders. In most cases, however, there was some aspect which could be considered similar in approach to the East High School program. Similarities of program are not totally absent in the two programs described in this section. But, synonymous grade level does seem to be the principal point of similarity between these programs and the program which is the subject of this paper.

Leamnson describes the initial results of the sophomore reading program in the Indianapolis Public Schools.[39] Reading laboratories were set up in each of the city's eight high schools. Each laboratory was equipped with twenty shadowscopes, a series of reading films, some workbooks, supplementary books and "reading materials of all kinds".[40] Time spent in the reading laboratory was taken from English class. Not all schools operated in the same

---

[39]George F. Leamnson, "Indianapolis Produces Better Readers," School Executive 74:64-67, December, 1954.

[40]Ibid., p. 65.

manner. The average class took two days per week for reading center work. The minimum was one day per week.

Time in the laboratory was divided among Iowa High School Reading Training Films, pacers, instruction in comprehension and word recognition, and discussion designed to increase wide reading.

According to standardized test results, students in the laboratory group had a 15% increase in rate after four months and a 6% increase in comprehension. Those who had no laboratory training had a 3% increase in rate and a 5% increase in comprehension during the same interval.

Beginning and end scores are not given and there is no statistical analysis of the possible meaning of the changes found. It is too bad that Leamnson did not wait until more reliable and meaningful results were available before publishing a report of the program's accomplishments. This article is included in several bibliographies of secondary reading programs recommended for study. The quality of the results and descriptions in the article hardly justify this prestige.

Marquis[41] describes a tenth-grade program for ten

---

[41]Bettylee F. Marquis, "Developmental Reading--New Albany High School," Journal of Developmental Reading 7:58-62, Autumn, 1963.

academic classes. This program used a long-term approach also. Each class spent one day per week in the reading laboratory for the entire year. The administration of the program is not fully explained; but machines were apparently used extensively, with some emphasis being placed on study skills. It should be noted that the program was tried with one section of general curriculum students as an experiment and it was concluded that the results did not justify continuing the program with this category of students.

In considering the results for the academic group, it should be noted that the upper students for this school were quite low in relation to the norm. On pre tests the mean reading rate was 178 WPM with 66% comprehension and mean rank was at the 34th percentile on the Cooperative English Test. At the end of the year the mean rate was 320 WPM with 81% comprehension. On the Cooperative English Test, the mean rank at the end of the year was at the 44th percentile. On the Cooperative English Test tenth grade norms were applied for the pre test and eleventh grade norms for the test at the end of the year. The mean gain in percentile rank did, therefore, truly and conservatively represent gain in skills.

Part IV--Summary and Conclusions

Only the best reported studies have been included in this review. Despite an attempt to demand relevancy and positive information as requirements for inclusion in this review, most of the studies must be declared highly unsatisfactory. A control group is rarely found; and, where it is found, the selection of control and experimental subjects in not always carefully accomplished.

A review of research on secondary reading programs leads to three generalizations. The best sounding programs are those where there is no control group and the program is relatively short. Programs which employ a control group in analyzing results have less impressive sounding outcomes. Least impressive sounding programs are those which are studied in progress and through a long follow-up with a careful comparison of experimental and control groups.

These trends, clearly evident in the research reported here, and the experience in Rochester[42] where one group made tremendous progress though its only "instruction" was pre and post testing, seem to suggest that the most important components of a successful reading program may be an optimistic instructor and an awareness of reading importance

---

[42]Barry and Smith, op. cit., p. 410.

by the students.

It should be added that the best sounding programs of all, not included in this review, are those where the evaluation is based entirely upon instructors' recollections of student comments.

Many of the components of the reading skill, especially those of an exterior, physical nature, have been carefully and scientifically analyzed. It does not seem unfair to state that reading programs have not generally been carefully analyzed to evaluate their effectiveness and that this is especially true at the senior high school level.

Possibly this lack of careful analysis is found at the secondary level because secondary reading is still a groping field. The literature is loaded with statements by authorities on what should be taught to senior high school students. It is noticeably less burdened with those who have specific, positive, proven approaches to the development of the delicately narrow and complex skills of the higher levels of total reading skill.

If this reviewer seems unfairly severe, it should be hastily added that the present research suffers some of the same shortcomings as those found in several of the programs reported here.

It is by no means claimed that the present research is

fundamentally different or superior to the majority of the studies reviewed. Those which have control groups are structurally more valid in their conclusions than the present study. However, even with control groups, short-term studies can be misleading in the credit claimed for a given approach to reading instruction.

In educational problems, we are often forced to assume that ultimate goals of learning are being achieved on the basis of fragmentary or short-term studies. Follow-up studies of reading programs, including the present study, indicate that, for those who really want to be certain about the value of a given instructional approach, short-term studies provide a hazardous avenue to evaluation in the field of secondary reading. Conversely, it appears that one who is only interested in selling an idea or approach can quite easily prove, through a short-term study, that the idea or approach is sound and successful.

CHAPTER III

THE PROGRAM STUDIED

Descriptions of the program taught to the sophomores of 1962-1963 at East High School, Duluth, Minnesota, are available in two published sources.[1] These sources also describe the type of student found at East High School during the period covered by this study. The following excerpts from these sources give a concise summary of the methods and goals of the reading program under consideration.

> The school enrolls approximately twelve hundred students in grades ten, eleven, and twelve and is located in the high socio-economic residential area of the city. Some idea of the above average nature of the student body can be gained by examining the intellectual ability, reading ability, and college matriculation characteristics of the group. Data from a group intelligence test indicates that the mean verbal and nonverbal intelligence quotients for over four hundred sophomores were greater than one standard deviation beyond the mean of the norm group for the test. In terms of reading ability, the average scores for rate, vocabulary and total comprehension fell above the 70th percentile on a standardized reading test. A recent follow-up study of East graduates showed that 80 to 85 per cent of the senior class enrolled in some institution of higher learning upon graduation. In comparison, it is estimated that 30 to 40 per cent of the graduates in the average high school enroll in an institution of

---

[1]Edward G. Summers, "Evaluation of Reading Gains in a Secondary School Reading Laboratory," The Reading Teacher 17:255-59, January, 1964; and, East High Reading Center Report. Duluth, Minnesota; Duluth Public Schools, November, 1962.

higher learning upon graduation.[2]

Although Summers comments further upon the atypical characteristics of students at East High School, the sense of his remarks is faithfully encompassed by the above quotation. The extremely long, uninterrupted quote which follows is primarily devoted to program description.

> . . .The program developed included the following general objectives:
>
> 1. to increase general reading speed and efficiency
> 2. to sharpen and extend comprehension
> 3. to improve study skills
> 4. to develop a flexible approach to reading
> 5. to develop general vocabulary and vocabulary specific to content areas
> 6. to extend interest and enjoy getting meaning from the printed page
>
> Organization of the Program
>
> The organization and evaluation of one six-week block of instruction for a group of five sophomore classes during the 1962-1963 school year is described in this study. The students were scheduled in the reading center for six weeks of time taken from their regular English class. The English teachers and the reading specialist provided the reading instruction. The students and teachers were oriented to the nature of the reading program through the use of a student handbook and teacher manual for the handbook. Prior to beginning the reading instruction the teachers discussed the handbook with their classes to develop readiness for the program. The handbook included objectives for the program, an explanation of the reading process and background material on the development of speed and efficiency, vocabulary building, increasing comprehension in various types of materials, and the improvement of of study habits. The materials to be used in the reading laboratory were also described.

---

[2]Summers, The Reading Teacher p. 255.

Although the general reading ability of the sophomore class was well above average, the five English classes which came to the reading laboratory for the period under study exhibited a wide range of ability on pretest scores. A range of scores from the 19th to the 99th percentile on rate, vocabulary, and total comprehension was common in all classes. A similar range existed in intellectual ability. Except for one enriched class, the sophomore English classes at East are not grouped for instruction. Thus, the reading laboratory was faced with the problem of developing a program for heterogeneous classes and individualizing instruction and materials sufficiently to meet the needs of the diverse students found within any one classroom. This is, of course, not a unique problem in education. However, descriptions of techniques for meeting this problem on the secondary level are not numerous, and evaluation, in statistical terms, of the effectiveness of such techniques is rare.

Description of the Reading Program

The instructional procedure for all five classes was essentially the same. A typical pattern of instruction for one class was as follows. The class was divided arbitrarily into two groups the first day. This was purely for scheduling purposes and was not an attempt to divide the group into homogeneous instructional units. Each instructional period of fifty minutes was divided into two sessions of twenty-five minutes each. This made it possible for two activities to take place simultaneously, and allowed the reading specialist to work with a small group of twelve to sixteen students while the English teacher supervised the activities of the other groups. Two activities were scheduled for each period, and the students switched activities at the halfway point of the period. Each student had a progress folder containing test scores, work sheets, and printed materials of various types, which was filed in the reading laboratory. Three major activities were scheduled: speed development, group work, and individual practice.

Speed development. Eighteen Shadowscope pacers were available in individual practice booths. Students began reading at the approximate rate indicated by pretest results. They were encouraged to slowly increase their rate at comfortable increments. To provide some transfer, pages were read alternately,

one on the pacer and one off the pacer. As reading material, students selected novels of their choice from a large number available in the laboratory, judged to provide a wide range of difficulty and interest level. A general check was made on comprehension by requiring the students to write a brief report on the material read. A minimum of two novels was read, with many students reading as high as five or six books during the period of reading training.

Group work. During this period, instruction was given on topics which were felt to be needed by the total group of sophomore students. Practice in developing flexibility in reading various types of material for different purposes was given. Vocabulary in general, and vocabulary specific to each of the content areas, was stressed. Study skills received considerable attention and included discussion of note taking, writing examinations, scheduling time, studying in various content areas, and developing discipline and pride in academic accomplishment. SQ3R was systematically taught as a study technique. Spelling demons of high frequency were analyzed and taught by use of experimental programmed learning materials. A wide variety of teacher-made and published materials on a number of difficulty levels were used to provide practice on skills taught.

Individual practice. The emphasis during this period was on the individual student's reading weaknesses. Based on standardized test prescores, informal analyses, and the judgement of the reading specialist and the English teacher, each student was placed in practice material which would help him develop areas where he was weakest depending on his ability and expectancy. During any one class period, when individual practice was scheduled, students might be found working on a variety of materials ranging from fourth grade to college level in difficulty. It would be impractical to list all the materials utilized. Most of the major materials developed within recent years were available to the reading laboratory for this phase of the program. As much as possible, the 3 to 5 per cent of the sophomores needing corrective and remedial work were helped individually during this period.

A typical schedule of activities for one week is shown below.

Mon.	Tues.	Wed.	Thurs.	Fri.
Group Work	Speed Dev.	Group Work	Group Work	Interest Reading
Speed Dev.	Indiv. Work	Speed Dev.	Indiv. Work	Indiv. Work

For the total six weeks in the reading laboratory, allowing for testing, each students received sixteen hours[3] of instruction and practice in each of the following: speed development, individual practice, and group work.[4]

Apparently the program used for these students varied slightly from the program for earlier classes described in East High Reading Center Report. A description of the reading program from this source follows.

> . .Basically, it consists of twenty-eight hours of developmental reading including shadowscope pacing, recreational and interest reading, group timing exercises using materials describing the reading process and skills needed in various subject areas, controlled reading filmstrips with comprehension questions, and workbooks and teacher made materials at various levels geared to each students (sic) reading needs as measured by a preliminary survey test. In addition, two sessions were devoted to material on how to read a textbook effectively (SQ3R) and an analysis of listening skills.
>
> Except for two late afternoon classes who (sic) missed several practice sessions because of athletic events, all classes received the same amounts of instruction. Testing occupied three hours, and controlled reading, shadowscope pacing, group timing and individual practice work all about six hours each. Each of the five classes were (sic) split into two groups and a schedule arranged to alternate two of the reading

---

[3]There is an error in either reporting or calculation here. The figures given would total forty-eight hours of work, while the absolute maximum would be thirty hours including any time used for testing.

[4]Summers, The Reading Teacher pp. 256-7.

activities during each period. A weekly schedule would appear as follows.

Monday	Tuesday	Wednesday
Controlled Reader	Group Timing	Controlled Reader
Individual Work	Shadowscope	Individual Work
Thursday	Friday	
Group Timing	Individual Work	
Shadowscope	Shadowscope	

The Population

The sample consisted of 146 sophomore boys and girls with complete pre and post data, scheduled in five English classes, who attended the reading center for a total of six weeks. All classes were taught essentially the same reading program with adjustments of a minor nature in the curriculum from class to class to adjust to individual differences in reading needs within the classes.[5]

The results of the program Summers describes are summarized in the tables on page forty-two.[6] Together, the results shown in these tables include pre and post test scores for ten of the fifteen classes included in the present study.

In another table, not reproduced here, Summers showed that all differences found in Table II were significant at the .01 level of confidence. (For purposes of clarity, parenthetical material has been added to the tables by the present author.)

---

[5]Summers, East High Reading Center Report p. 4.

[6]Table I is taken from The Reading Teacher p. 258. Table II is taken from East High Reading Center Report p. 6.

TABLE I

Significance of the Difference Between Pre and Post Test Scores for all Groups for the Diagnostic Reading Tests--Survey Section

Group	N	Speed*			Vocabulary			Comprehension		
		Pre Mean	Post Mean	t(1)	Pre Mean	Post Mean	t	Pre Mean	Post Mean	t
Total	151	271	423	23.16	41.30	46.57	10.07	11.58	14.99	12.18
High Ability	51	288	477	12.25	47.33	50.96	4.19	13.71	16.49	7.91
Middle Ability	49	273	409	12.47	42.80	47.33	7.98	11.59	14.47	5.74
Low Ability	51	252	383	10.78	33.84	41.16	8.27	9.43	13.98	8.15

*The speed score is expressed in words read per minute, the vocabulary score is the number correct out of sixty items, and the comprehension score is the number correct out of twenty questions.

(1)(All t scores were found to be significant at the .01 level.)

TABLE II*

Pre and post test mean differences and means for the total sample on rate, total reading and rate of comprehension scores. (n=146)

	pre test mean	post mean	diff.
Rate (WPM)	270.91	444.89	173.98
Total RDG. (100 possible)	66.58	69.97	3.39
Rate of Comp.	212.08	262.09	50.01

*This table is reproduced in the style in which it was originally presented.

In East High Reading Center Report Summers did not report the results for the following factors included in the present study: Comprehension at measured rate of reading, vocabulary, and total comprehension. In the Reading Teacher article Summers did not report the following factors covered by this study: Total comprehension, total reading, and rate of comprehension.

The impact of these omissions upon evaluation of the program will be discussed in chapter six.

Materials used in the reading center, but not named in these reports, included: Reading for Meaning, Strang's Study Type of Reading Exercises, Reading for Understanding, and an S.R.A. Reading Laboratory.[7]

Further description of the reading program which is the subject of this study would be possible. However, the reliability of that description would depend upon the long-term memory of others than the reading specialist. Since the present study evaluates the efficacy of the program taught and reported by Summers, propriety suggests that the present evaluations be based upon Summers' descriptions, even though those descriptions were somewhat limited by the

---

[7]The use of these materials is known through post-program evidence of use, student answer sheets etc., and through conversation with English teachers who worked with the program.

problems involved in publication.[8]

This study reports the results of 1962-1963 pre and post tests for all students still enrolled at East High school in December of 1964 who took the test without observable errors in following directions on the three occasions for testing. A number of students left East High School during the interval between the pre and follow-up tests. In addition, quite a few students had, on one or more of the three tests, failed to record their rate score properly. These students were eliminated from the study. 413 students took the first test; 343 of these students are included in the present study.

This has been a broad and summarized view of the students and program evaluated in this study. The next chapter is devoted to the approaches used to accomplish this evaluation.

---

[8]Summers, The Reading Teacher p. 257.

CHAPTER IV

DESCRIPTION OF THE STUDY

This chapter will explain the scores used in evaluating the reading program described in chapter three. In addition, the rationale for the division into intelligence groups and the methods of evaluation will be described. Evaluation of the program was approached from as many directions as the available data appeared to permit.

A multiple approach to evaluation was undertaken because of the lack of a control group for this study. There is no possibility of providing statistical proof to support any conclusion found in this study.

The evidence provided by these approaches is meant to supply as reasonable a basis as possible from which to draw what must remain intrinsically subjective conclusions.

Test Administration and Scores

The pre test upon which the present evaluation is based was administered to all sophomores attending East High School in September of 1962. The post test was administered at the close of reading center instruction. By the end of January all sophomore English classes had completed their sessions in the reading center. Approximately one-third

took the post test in October, approximately one-third took the post test in November, and approximately one-third took the post test in January, 1963. The follow-up test was administered to all students included in this study in December of 1964.

The pre test and the follow-up test were identical--Form B of the Diagnostic Reading Tests, Survey Section: Upper Level. Form A of the same test was administered as a post test. These forms are considered so comparable in difficulty that standardization tables for both are based on Form A.

Pre and post tests were administered by Summers. The follow-up test was administered by the author.

These tests yield scores in rate (words per minute, often abbreviated WPM in this paper), comprehension of story-type reading, vocabulary, total comprehension, and total reading. In addition to these scores, it is possible to derive scores in comprehension of study-type materials and rate of comprehension. The tests have been standardized for rate, vocabulary, total comprehension and total reading.

The rate score is the average number of words read per minute for three minutes, without reference to the student's degree of comprehension.

Rate of comprehension, often abbreviated R/C in this

paper, might be called an index of net reading. It is derived by multiplying the per cent correct on the comprehension check of story-type reading by the rate at which that material was read. A table in which rate was rounded to the nearest ten before determining R/C was used in finding this score for each student. The true R/C could, therefore, vary by as much as five points from the approximations found in the present study. Since a five point variation would be of very minor importance, R/C is reported as an exact score for purposes of convenience.

When R/C is discussed, it must be remembered that this refers only to story reading. There is no score which measures the reading rate of these students in study-type reading.

Comprehension of story-type material is the number correct out of twenty questions on the story reading. This is the same material which is used for the timed reading which determines the rate score.

The study-type comprehension score is derived by subtracting the story comprehension score from the total comprehension score. The total possible on this score is twenty.

Vocabulary score is the number correct out of a possible sixty correct responses on multiple choice questions.

The students are allowed ten minutes for this section of the test.

The total comprehension score is the number correct on story-type reading plus the number correct on study-type reading. There are a total of forty possible correct responses on the two sub-tests which constitute this score.

Total reading score is the sum of the total comprehension and vocabulary scores. The number of possible correct responses on this total of all test questions is one hundred.

Groups and Intelligence

In addition to test scores, intelligence scores for the students included in this study are utilized. These scores, taken directly from school records, are for the verbal section of the Lorge-Thorndike intelligence test. This test was administered to these students during the same semester that they participated in the reading program under study.

Verbal intelligence only is considered in this study because of its close relationship to reading ability. When the terms intelligence or I.Q. are used in chapter five, it is assumed, for the purposes of this paper, that this verbal score fulfills the meaning of those terms.

In some cases, students exceeded the limits of the Lorge-Thorndike test and scores were, therefore, recorded as 150 plus. In such instances scores were transcribed as 150, and they appear as 150 within the statistics of this paper.

The results of testing are reported for the group as a whole, but the results are also broken down into intelligence levels. It was felt that a concentrated period of reading emphasis, especially one where speed was a prominent factor, might affect students of differing intelligence levels in different ways. For this reason, students were divided into groups of narrower intelligence range.

There was no attempt to divide the students into three equal groups. The attempt was to approximate an East High School intelligence curve rather than a normal intelligence curve. Students with I.Q. 130-150 were placed in the group designated High. Students with I.Q. 110-129 were placed in the group designated Middle. Students with I.Q. 84-109 were placed in the group designated Low. When all of the students in the study are being considered together this is designated the Whole group. Whole refers only to the entire number of students in the study considered as one group. Total is used to refer to the sum of students or scores within any group.

Mean Scores

The mean scores found at each point of testing are the primary means of evaluation utilized. A study of the statistical significance of the difference of the means would be inappropriate for two reasons.

Since there is no control group, there is no statistically sound means of projecting the results found here to any other population as a prediction of program effectiveness with that population. Secondly, there would be no point in attempting projections of reliability of the scores for this group when all scores are included in the mean comparisons.

Mean score changes are compared for the High, Middle, Low, and Whole groups.

Percentile Rank

The percentile ranks of the mean scores for those scores where a norm has been established are reported for the High, Middle, and Low groups.

Almost always this percentile rank is reported in terms of the nearest obtainable percentile rank. This element of approximation is necessary because the mean score is not often one which can actually be obtained by any one student on the test. It is impossible, for example, for a single

student to obtain a fractional number correct on the test.

Expectancy

Changes in reading ability from tenth to twelfth grade are also evaluated on the basis of expected change. This is accomplished by noting the mean percentile rank on a given test as sophomores and assuming that the mean percentile rank as seniors should be the same. The expected mean raw score for twelfth grade, then, is the one found on twelfth grade norms for the percentile rank the group attained as sophomores.

These expectancies and comparisons with actual twelfth grade scores are reported on the basis of both pre and post test scores.

This method of evaluation presents many hazards. These are not average students in an average high school setting. How their reading skills would develop over this period of time without a reading program cannot be definitely known. This is only another attempt to find a clue to program effectiveness.

Correlations

As will be seen in chapter five, none of the preceding methods of evaluation, all basically derivations of mean score comparisons, proved to be a very conclusive method

of evaluating the reading program. As an afterthought, changes in reading skills were studied by correlations within small random samples of the High, Middle and Low groups.

A correlation study of this nature can give no absolute answers. However, it was felt that the degree of relationship between reading changes occurring during the reading program and reading changes occurring later might provide a clue to the program's follow-through value.

Since the validity of projecting the correlation results to the total of the groups they represent is open to question, the results and discussion of the correlation studies are presented separately in the appendix of this paper.

Summary

Three major means of evaluation were attempted--comparison of mean scores, percentile rank, and expected gains. None of these can prove anything about the reading program under study. They do, however, provide as many avenues of evaluation as are reasonably open to the author. The primary value of this study lies in the time elapsed from the beginning to the end of the period studied, along with the multiple approach to evaluation. The lack of a control group is a regrettable limiting factor.

CHAPTER V

RESULTS OF THE STUDY

This chapter will attempt to evaluate the reading program discussed through the use of measures and comparisons described in chapter four. The conclusions which various measures suggest are not always consistent with one another. In so far as possible, the results of each approach to evaluation will be analyzed separately. Overall conclusions about the value of the reading program will be reserved for chapter six.

Clarifying comment upon the value and meaning of various measures and scores is an essential ingredient of this section of this paper.

The first part of this chapter will be devoted to a comparison of mean scores found at each of the three testing points--1962-1963 previous to reading training, 1962-1963 following reading training, and December of 1964.

Part I--Mean Scores

Whole Group

The pre, post, and follow-up mean test results for the 343 students who constitute the Whole group are shown in Table III, below.

TABLE III
MEAN PRE, POST, AND FOLLOW-UP READING TEST SCORES OF ALL STUDENTS (WHOLE GROUP) (n=343)

Skill	Pre test	Post test	Follow-up test
Rate in WPM	280.2	461.2	339.0
Rate of Comp.	216.4	287.6	255.8
Story Comp.	15.3	12.5	15.0
Study Comp.	12.4	15.0	16.6
Vocabulary	42.0	47.3	49.7
Total Comp.	27.7	27.5	31.6
Total Rding.	69.7	74.8	81.3

Rate and rate of comprehension

Reading rate measured without reference to comprehension carries little meaning. Table III, above, does show, however, a great increase in rate from pre to post test, a great decrease from post to follow-up test, and a modest increase from pre to follow-up testing.

Rate of comprehension (WPM multiplied by per cent correct on comprehension check) yields a more meaningful score. The R/C shows a mean increase of 71.2 points or 32.9% from pre to post test and a mean increase of 39.4 or 18.2% from pre to follow-up test. The decrease from post to follow-up test was 31.8 points. Obviously, the rate of comprehension scores show both meaningful gains and meaningful loss.

Interpretation of the R/C scores has one limiting factor. There is no correction factor for guessing in the test administered. For this reason, a student scanning rather than actually reading the material may well attain a higher R/C score than a student reading at a more reasonable rate and comprehending more fully. In other words, the laws of chance involved in a multiple choice test favor a higher R/C for a greater rate more than would be the case were a correction factor used. Therefore, the higher the per cent correct on the comprehension check, the greater

the validity of the R/C score as a measure of net reading.

As an example of this, let us suppose that a student scores sixteen correct out of twenty questions at a reading rate of 200 WPM on a pre test. Since there are four possible choices for each question, a correction factor would ordinarily reduce this score to 15. R/C without the correction factor (80% X 200) would be 160. With a correction factor R/C would be (75% X 200) or 150. Assume now that this student scores eight correct out of twenty at a rate of 600 WPM on a post test. Without a correction factor, the student's R/C score would be considerably higher than on the pre test, (40% X 600) or 240. A correction factor would reduce the score from eight to five and the per cent correct from 40% to 25%. Following this procedure the R/C would be (25% X 600) or 150--exactly the same as for the pre test.

The differences in mean per cent correct on the comprehension check of timed reading for the three testing points are not, of course, nearly so extreme as those given in the example. The example is only meant to demonstrate that, because of the higher mean per cent correct on the comprehension check (story comprehension, Table III, page 54), the pre test R/C has the greatest validity as a measure of net reading. Because of the lower per cent

of correct answers on the comprehension check for the post test, the post test R/C has the least validity as a measure of net reading. Following the same line of reasoning, the follow-up test R/C is almost, but not quite, as valid as the pre test R/C.

Story comprehension in timed reading

As already noted in the discussion of R/C validity, there was a considerable drop in comprehension in the timed reading section of the test from pre to post test.

The mean scores shown for this sub-test in Table III, page 54, indicate the number correct out of twenty. The drop from a mean of 15.3 to 12.5 from pre to post test indicates a drop in comprehension from 76.5% to 62.5%, a considerable loss. Although the difference between pre and follow-up test mean score is negligible, 1.5%, it should be noted that not all of this comprehension loss was regained, according to the results of the follow-up test.

Although temporary comprehension losses are sometimes expected, especially for some compulsive readers, at the beginning of rate improvement training, most authorities would question the value of increasing reading rate at the expense of significant comprehension losses.

Comprehension in study-type materials

This score is not separated from the previously discussed comprehension score in scoring or standardization of the test. However, it is timed separately and administered as a separate sub-test.

The scores given in Table III, page 54, indicate the mean correct out of a possible twenty. On this test the student reads four series of paragraphs. Each series is followed by five comprehension questions. The student is allowed fifteen minutes to complete this section of the test.

Observation by this author would indicate that speed is a factor for only the slowest readers. The vast majority of East High School tenth graders have no difficulty in finishing this part of the test prior to any special reading training.

Table III indicates substantial gains in study-type comprehension from pre to post test. It increased 2.6 points or 21% from pre to post test. A further gain of 1.6 points or 10.7% above the post test mean can be noted for the follow-up test.

It seems fair to say that, on the basis of mean comparisons, the gain shown in this type of comprehension is one of the most meaningful and unambiguous positive results

of the reading program indicated by the mean results of these tests.

Vocabulary

The vocabulary improvement shown from pre to post test is dramatic on the surface and it deserves close examination. Table III, page 54, shows a mean gain of 5.3 points or 12.5% from pre to post test. The nature of this improvement may be better explained by an increase in reading speed contributing to score rather than by true increase in vocabulary.

Attributing vocabulary gain to speed is based upon five factors.

The present author, who has administered this test to over eight hundred sophomores before and after reading rate training, has noted that many fewer responses are left blank at the end of this sub-test after rate emphasis, even though no emphasis has been given to vocabulary improvement.

Secondly, it is illogical that even the finest vocabulary instruction could cause a true increase of 12.5% in the vocabulary of a superior group of tenth graders in six-weeks' time. If vocabulary were truly being measured, this would indicate an increase of no less than 625 words. (These figures are based on the conservative assumption

that the average sophomore can recognize the meaning of five thousand words beyond those most basic to the language.) An increase of 625 words would amount to learning and retaining more than 21 words per laboratory session!

Since this is an above-average group in vocabulary, the gross increase would almost certainly have to exceed the hypothetical figures given above to account for a 12.5% increase.

A third factor contributing to the conclusion that the vocabulary score increase does not represent true vocabulary increase is shown in Table IV, page 61. The mean increase in vocabulary raw score and the per cent of mean increase in vocabulary from pre to post test are both inversely related to intelligence.

Because the High and Middle groups have a more rapid reading rate on the pre test than does the Low group, it is sensible to conclude that the improvement in speed would have less effect upon mean vocabulary scores for those groups than it would have upon the mean vocabulary scores of the Low group.

A temporizing factor in this judgement, however, is the proximity of test ceiling. The Low group mean was slightly under two-thirds of a perfect score on the post

TABLE IV
RELATIONSHIP OF INTELLIGENCE
TO VOCABULARY GAINS

	High Group	Middle Group	Low Group
Pre vocabulary	50.5	43.3	33.0
Post vocabulary	54.2	48.5	39.6
Change	3.7	5.2	6.6
Per cent of change	8.0	12.0	20.0
Mean I. Q.	137.3	119.3	101.4

test, while the Middle and High group means both exceeded that point on the pre test. The influence of test ceiling upon scores for the various groups is discussed further in the section of this chapter dealing with comparisons among intelligence groups.

Despite the complication in interpretation of these scores caused by the problem of test ceiling, it seems relevant to note that it is contrary to educational research experience to find the less intelligent and less proficient members of a group progressing more rapidly in a given skill than the more proficient and more intelligent members of the same group, when all are receiving the same amounts and kinds of instruction.

A fourth factor in discounting the pre to post vocabulary gain as a true measure of vocabulary improvement is found in comparing pre to post gain with post to follow-up gain. These students, who had a mean gain of five points from pre to post test, only managed an additional gain of 2.4 points during the two-year interval between the post and follow-up tests. It should be remembered that this interval includes two-thirds of these students' senior high school experience.

The final evidence of speed as a primary factor in vocabulary gains is found in the test and test manual

themselves. The vocabulary sub-test demands sixty responses in ten minutes--an average of one response every ten seconds if the student is to complete the test. The first, thirtieth, and sixtieth vocabulary questions from Form A of the test administered follow as examples of the difficulty level of the task of finishing this sub-test in the time allowed. (The vocabulary sub-test is found in the middle of the test administered.)

21. Group of persons united for some purpose
(1) clan (2) organization (3) brotherhood
(4) partnership (5) religion

50. An expression consisting of two terms
(1) equation (2) binomial (3) bilateral
(4) dual (5) biannual

80. A band of gold, jewels, or flowers worn around the head as an ornament
(1) halbert (2) frieze (3) demijohn
(4) replica (5) tiara[1]

Probably the most valid argument for attributing at least part of the gain in vocabulary score to gain in speed is based upon the directions for interpreting the vocabulary test. Those directions follow in their entirety. (Underlining is the present author's.)

> Vocabulary: Score 2--This measure of familiarity with the meanings of words is sometimes influenced by extremely slow reading. In the time allowed them, such

---

[1]Committee on Diagnostic Reading Tests, Inc., Diagnostic Reading Tests Survey Section: Upper Level pp. 12, 14, 16. Published by the Committee on Diagnostic Reading Tests, Inc., Mountain Home, North Carolina, 1947

> students may complete too little of the test for adequate measurement of their meaning vocabulary. Among college freshmen, for example, attempting fewer than 40 items implies a marked handicap, due probably to the rate at which items were read. Permitting the student to attempt the remaining items, without time limit, may help. If the untimed score is 20 to 30 percentiles higher than the original timed score, one may conclude that the student's breadth of vocabulary is not as limited as his first trial indicated. While not directly comparable with the timed scores of other students, the untimed score of the slow reader gives a better estimate of his vocabulary potential than the score from a timed vocabulary test.[2]

It should be noted that the mean increase in vocabulary score is not meaningless, even when its significance as a measure of vocabulary increase is discounted. Although not definable, this score gain may be one of the most reliable indicators that increase in speed of comprehension actually did take place during the reading program.

Although the mean increase in vocabulary score from pre to post test cannot reasonably be entirely accounted for by actual vocabulary growth, some growth did take place. It is the contention of this paper that the definition of that growth cannot reasonably be encompassed by the term vocabulary.

---

[2]Committee on Diagnostic Reading Tests, Inc., Diagnostic Reading Tests, Survey Section, Directions for Administering p. 9. Published by the Committee on Diagnostic Reading Tests, Inc., Mountain Home, North Carolina, 1947.

Total comprehension and total reading

Both of these scores are composites of mean scores already discussed. Total comprehension is a combination of the study-type and story comprehension questions. Total reading is a combination of vocabulary and total comprehension scores.

Total comprehension was down .2 of a point from pre to post test and total reading was up 5.1 points. All of the gain in total reading, plus .2 of a point, is attributable to the vocabulary increase already discussed.

From post to follow-up test, total comprehension showed a mean gain of 4.1 points. For total reading score the gain from post to follow-up test was from a mean of 74.8 to 81.3.

Whole group conclusions

Mean scores for the Whole group indicate the following general results for the program.

Rate was increased significantly at the expense of a significant loss in comprehension as shown by rate and story comprehension scores. A small, and probably not very meaningful part of this comprehension was not regained on the follow-up test two years later. From the post test to the follow-up test mean rate dropped 122.2 WPM. The change from pre test to follow-up test was a net gain of 58.8 WPM.

Study-type comprehension was increased meaningfully from pre to post test and a further gain of smaller, though meaningful magnitude, was found between post and follow-up test.

The gain in vocabulary from pre to post test must be viewed with suspicion as a measure of vocabulary, since it surpasses credibility. At least in part, the vocabulary gain may be considered in some way a measure of undefinable speed increase. The vocabulary gain from post to follow-up test, while small, appears credible and meaningful.

Rate of comprehension gain from pre to post test, while not totally reflecting valid gain, can be said to indicate substantial gain in net reading from pre to post test and a less spectacular gain from pre to follow-up test.

Comparisons Based Upon Different Levels of Intelligence

It may be noted from quotations of Summers in chapter three and statements by the present author in chapter four that the students of East High School do not approach a normal curve in intellectual ability. The mean intelligence of each of the three groups, shown in Table IV, page 61, indicates that the designations High, Middle, and Low for these groups are convenient rather than descriptive terms. The Low group members would, for the most part, fall within the limits of an average group in many secondary school

settings.

The rationale for comparing differing intellectual levels, as noted in chapter four, was the concern that a concentrated program of the kind studied might influence the reading skills of students of these differing intelligence levels in different ways.

Table V, page 68, gives the number and pre, post, and follow-up test score means for each group. When using Table V as an aid to understanding the text, it may be helpful to note that scores for the various skills are separated by a horizontal line. Where more than one sub-test is discussed in a section these scores are found consecutively in the table.

Repetition of comments on the meaning of individual sub-test scores, which have been discussed for the Whole group, will be minimal in discussing scores of the varying groups.

Rate and rate of comprehension

The facts and figures discussed here are either found in or may be deduced from Table V, lines 1-6, page 68.

From pre to post test the mean gain in rate was consistent with intelligence. The High group had the greatest mean increase in rate and the Low group had the lowest mean increase. The per cent of increase was also higher for the

TABLE V

MEAN TEST SCORES OF HIGH, MIDDLE, AND LOW GROUPS

Skill	High group (n=70)	Middle group (n=182)	Low group (n=91)
RATE			
(1) Pre	327.2	278.2	248.2
(2) Post	558.5	456.1	396.5
(3) Follow-up	425.0	328.6	293.6
RATE OF COMPREHENSION			
(4) Pre	272.5	224.1	157.8
(5) Post	398.4	281.3	214.9
(6) Follow-up	354.5	254.8	181.8
STORY COMPREHENSION			
(7) Pre	16.8	16.1	12.4
(8) Post	14.5	12.5	10.8
(9) Follow-up	16.7	15.5	12.5
STUDY COMPREHENSION			
(10) Pre	15.5	12.7	9.8
(11) Post	17.6	15.1	13.0
(12) Follow-up	18.8	17.1	14.4
VOCABULARY			
(13) Pre	50.5	43.3	33.0
(14) Post	54.2	48.5	39.6
(15) Follow-up	56.2	51.0	41.9
TOTAL COMPREHENSION			
(16) Pre	32.3	28.8	22.2
(17) Post	32.1	27.6	23.8
(18) Follow-up	35.5	32.6	26.9
TOTAL READING			
(19) Pre	82.7	72.1	55.2
(20) Post	86.4	76.1	63.4
(21) Follow-up	91.7	83.6	68.8

High group (71%), but the Middle and Low groups had identical per cents of increase in rate (60%).

Pre to post rate of comprehension changes among the groups were not as consistent with rate change as one might have expected them to be. The increase in R/C was considerably higher for the High group; but the increase in R/C for the Middle and Low groups was almost identical, 57 plus. The per cent of increase was highest for the High group at 46%; the Low group's increase, 36%, fell midway between the High and Middle groups in this respect. Middle group increase, of course, was 26%.

All groups measured a loss in both rate and rate of comprehension from post to follow-up test. The losses did not equal the gains found between pre and post tests for any of the groups.

Losses in rate were consistent with gains in the sense that the group which had gained the most, the High group, lost the most, while the Low group, which gained the least, lost the least. Being 24%, 28%, and 26% for the High, Middle, and Low groups respectively, the losses for the three groups were quite even on a per cent basis.

In rate of comprehension the loss and per cent of loss were both lowest for the Middle group. Although the High group experienced the greatest mean loss, its per cent of

loss was considerably lower than the Low group's.

Story and study-type comprehension

The facts and figures discussed here are found in or may be deduced from lines 7-12 of Table V, page 68.

Each of these kinds of comprehension is measured by twenty questions. The story comprehension score is the comprehension check for timed reading. Thus, this score reflects to some degree the emphasis students were placing upon speed in reading at various test points.

From pre to post test all groups experienced a drop in story comprehension. The per cents of correct answers on the pre test were 84%, 80.5%, and 62% for the High, Middle, and Low groups respectively. The groups ranked in the same order on the post test; but the gap between the High and Middle groups widened, while the gap between the Middle and Low groups narrowed. On the post test the per cents correct dropped to 72.5%, 62.5%, and 54% for the High, Middle, and Low groups respectively.

The mean per cent correct on study-type comprehension rose for all groups from pre to post test. On the surface, the figures suggest that improvement in study-type comprehension was inversely related to intelligence. Such a conclusion probably should be withheld. The mean of the High group on the post test was only 2.4 points from test

ceiling. Despite the larger gain in score from pre to post test for the Low group, the Low group's mean score was still almost three times as far from test ceiling.

On the follow-up test all groups gained over the post test on story comprehension. It should be noted, however, that the High and Middle groups did not quite regain the pre test level of comprehension, while the Low group mean in story comprehension on the follow-up test slightly exceeded pre test mean.

The problem of test ceiling as a limitation in measuring High group improvement is clearly shown in the post test scores of study-type comprehension. The Middle group showed the greatest gain from post to follow-up test, but this group was more than twice as far as the High group from test ceiling on the follow-up test. The Low group's gain was also greater than the High group's; but the mean score for this group was farther from test ceiling on the follow-up test than the High group's mean was at its lowest point, on the pre test.

Vocabulary

The weaknesses of considering the vocabulary test score as a true measure of vocabulary improvement between pre and post test have already been discussed at length in connection with Whole group scores. The vocabulary scores for

the different intelligence groups are reported in lines 13-15 of Table V, page 58. Further discussion of pre and post score changes at this point would be redundant.

Despite the much greater distance from test ceiling, the Low and Middle groups gained only slightly more than the High group from post to follow-up test. Between these points the change in per cent correct was about 4% for all groups.

Total comprehension and total reading

These scores are composites of the scores already discussed. Lines 16-21 of Table V, page 68, present a convenient summary of the total picture of reading changes other than in rate.

Total comprehension and total reading scores ranked consistently with intelligence at all points of testing. That is, the High group scores were always the highest and the Low group scores were always the lowest at a given testing point.

It is interesting to note that the Low group's best scores, on the follow-up test, did not equal the higher groups' scores at their lowest points. The Middle group's follow-up scores, however, did slightly exceed the pre total comprehension and total reading scores of the High group. On the surface, this fact would seem to indicate

that the progress of the Middle group over the period studied was more impressive than the progress of the Low group. It would probably be better to consider this only as an interesting speculation.

Conclusions on Mean Comparisons

Pre to post test mean differences suggest that all intelligence levels benefitted from the reading training given. All of the preceding statistics and all of those to follow do contain a built-in margin for error.

As noted in chapter four, all of the pre tests were given at the beginning of the school year. The post tests were administered at the end of the six-weeks in the reading center, but the period of training came immediately after the pre test for only approximately one-third of the students. Another third of the students received training during the interval 7-14 weeks after the pre test, and the rest of the students received training during the interval 13-18 weeks after the pre test.

It is reasonable to guess that a reading test administered after the practice effect of required school reading had taken place might yield higher mean scores than the same test given immediately following a three-month vacation period would yield, whether or not a special reading course had been taught in the interval.

The comparison of mean scores suggests that the greatest benefit to the students was in the area of study-type comprehension, although the measurement of this gain, at least for the High group, is probably somewhat hampered by test ceiling.

As a measure of some type of comprehension speed, the vocabulary sub-test showed a good deal of growth taking place, especially for the Middle and Low groups. Its significance as a measure of growth in recognition of word meanings, however, must be discounted.

Although rate improvements were spectacular from pre to post test for all groups, the value of this growth must be questioned both on the basis of a drop in comprehension and on the basis of rather spectacular loss from post to follow-up test.

Rate of comprehension showed improvement for all groups from pre to post test, a reflection of rate increase. Its increase is much less, however, than is the increase for the rate because this score includes the drop in comprehension as a factor. The R/C index probably deserves consideration as the true measure of rate improvement from pre to post test and rate regression from post to follow-up test, despite the previously discussed limitations on the validity of this score for this particular reading test.

Follow-up test means showed that the progress from pre to post test implied excessive credit for the value of the reading program, especially in the areas of vocabulary and rate growth. These areas appeared to benefit most on the basis of pre and post scores, but the long-range examination provided by this analysis suggests that the actual area of greatest permanent gain was probably in the previously mentioned study-type comprehension.

Part II--Evaluation of Percentile Rank at Testing Points

A rather thorough overall view has been presented for the Whole group in the mean comparisons section. All further evaluations have been undertaken only for the differing intelligence groups. Not all the areas involved in the analyses of mean scores are present for comparisons based upon percentile rank. Only rate, vocabulary, total comprehension, and total reading have been standardized.

As mentioned in chapter four, percentile ranks and changes in ranks are expressed in terms of the nearest obtainable percentile rank. The mean score is often one for which no exact percentile rank is given, since it contains fractional units.

Percentile rank comparisons can very easily distort the actual gain or loss shown by a given mean raw score

change. On the other hand, percentile comparisons can show how these students compare with a wider population where raw score changes cannot.

Pre and post test percentile ranks shown in Table VI, page 77, are for tenth grade norms. Follow-up percentile ranks are, of course, based upon twelfth grade norms.

Rate

Table VI, below, shows that gains in percentile rank on rate from pre to post test were worthwhile for all groups. Although there was considerable loss between post and follow-up test for all except the High group, the percentile rank on the post test in comparison with the pre test is still impressive. The overall percentile rank gains were 10 points, 16 points, and 20 points for the High, Middle and Low groups respectively.

Vocabulary

Percentile rank changes in vocabulary showed high volatility for the Middle and Low groups as previous discussion would suggest that they should if the vocabulary pre to post test gains were not truly accounted for by vocabulary improvement. The overall vocabulary percentile rank gains, pre to follow-up test, were modest. They were two, three, and seven percentile points for the High, Middle and Low groups respectively. The overall rise in vocabulary

TABLE VI
PERCENTILE RANK OF THE MEAN SCORES FOR DIFFERENT INTELLIGENCE GROUPS

Skill	High group (n=70)	Middle group (n=182)	Low group (n=91)
RATE			
Pre	88	69	53
Post	99	99	96
Follow-up	98	85	73
VOCABULARY			
Pre	94	79	48
Post	97	91	70
Follow-up	96	82	55
TOTAL COMPREHENSION			
Pre	90	75	37
Post	90	70	48
Follow-up	96	83	44
TOTAL READING			
Pre	94	79	42
Post	97	86	61
Follow-up	97	85	50

status appears reasonable and worthwhile.

Total comprehension

Since this measure included both story and study-type comprehension, pre to post test losses in story comprehension and gains in study-type comprehension are masked in percentile rank comparisons. Again it would seem proper to emphasize the long-range changes. From pre to follow-up test the percentile rank standings rose 6, 8, and 7 points for the High, Middle and Low groups respectively.

Total reading

Table VI, page 77, shows that all pre to follow-up percentile rank changes for total reading were also upward. Pre to post gains were not held for the Middle and Low groups, which had the greater gains from pre to post test. The High group, however, which moved only three percentile ranks from pre to post test, held at the very difficult to maintain ninety-seventh percentile rank standing.

Conclusions

Despite the wide differences among the groups in percentile rank change from pre to post test, pre to follow-up changes for all groups were within a very similar range. On rate the percentile rank gains ranged from 10 for the High group to 20 for the Low group from pre to follow-up test. On all other percentile rank scores of means the

range of gain was 2 to 8 percentile ranks from pre to follow-up test.

The desirability of the rate gains for the Middle and Low groups may be questioned. Eighty per cent comprehension is often mentioned as a desirable level for story material. This was not attained by either the Middle or Low groups on the follow-up test.

The percentile rank comparisons probably raise more questions than they answer. As previously mentioned, the pre test was administered in September of the tenth grade year. It seems possible that the modest pre to follow-up percentile rank gains, other than for rate, may be entirely accounted for by the fact that follow-up tests were administered in December rather than September of the twelfth grade year. The small change in raw score on the follow-up test which would be necessary to have percentile ranks for vocabulary, total comprehension, and total reading mean scores stand at the same point as on the pre test is demonstrated in the next part of this chapter.

Part III--Expected Twelfth Grade Level of Reading Skills

There are two purposes behind the expectancy approach to evaluation. One of these purposes is to place the

percentile rank changes discussed in the previous section in proper perspective. In some cases, the differences shown in percentile rank standings tend to create an impression of greater change having taken place than was actually the case.

A more important purpose of this section is to provide an estimate of the degree of growth one should reasonably expect these students to achieve, on the basis of tenth and twelfth grade reading test norms, in the interval between the original tests and the follow-up test.

As explained in chapter four, the expectancy level is based upon the supposition that students should normally maintain an identical percentile rank from tenth to twelfth grade. This supposition is, of course, open to question. It probably would be indefensible to attempt to base an expected level of skill upon percentile rank on an individual basis. The procedure does seem to have more validity when applied to the mean of a group. The averaging process in the group situation should ordinarily level-off the variation which is normal among individuals.

The expected level for twelfth grade, then, is found by taking the percentile rank achieved as sophomores, finding the identical percentile rank on twelfth grade norms and reading the standardization table in reverse to arrive

at the raw score for that percentile rank. The raw score found at that point then is considered the raw score which should be expected on twelfth grade tests.

It could be argued that staying in school through grade twelve, as all the students included in this study did, is in itself a selective process. Although this is true, it does not invalidate the percentile rank method of arriving at an expectancy. The norm group upon which twelfth grade percentile ranking is based has the same selective factors operating within it.

Rate

Table VII, page 82, shows that all groups exceeded the expected level in rate for twelfth grade on the basis of pre tests. In no group did the twelfth grade rate score attain the level which would be expected on the basis of the post test. It should be noted that tenth grade rate scores on the post test far exceeded the level necessary for the Middle and High groups to be ranked at the 99th percentile. For this reason the expected level found in Table VII is lower than the rate these groups actually attained as sophomores.

The excess of the actual rate score over the pre test expectancy follows in line with intelligence levels, the High group exceeding expected level by the widest margin

TABLE VII
TWELFTH GRADE MEAN RAW SCORES EXPECTED ON THE BASIS OF PRE AND POST TEST PERCENTILE RANKS, ALONG WITH ACTUAL TWELFTH GRADE MEAN RAW SCORES

Expected and acutal scores	High group (n=70)	Middle group (n=182)	Low group (n=91)
RATE			
Expected 12th grade score based on pre test %ile rank	330	285	265
Expected 12th grade score based on post test %ile rank	455	455	395
Actual 12th grade score	425	328.6	293.6
VOCABULARY			
Expected 12th grade score based on pre test %ile rank	55	50	40
Expected 12th grade score based on post test %ile rank	57	54	47
Actual 12th grade score	56.2	51	41.9
TOTAL COMPREHENSION			
Expected 12th grade score based on pre test %ile rank	34	32	26
Expected 12th grade score based on post test %ile rank	34	31	28
Actual 12th grade score	35.5	32.6	26.9
TOTAL READING			
Expected 12th grade score based on pre test %ile rank	89	81	66
Expected 12th grade score based on post test %ile rank	92	85	73
Actual 12th grade score	91.7	83.6	68.8

and the Low group exceeding expected level by the least amount.

For all groups the excess over pre test expectancy seems large enough to be considered meaningful.

Vocabulary

As was seen before for rate, Table VII shows that all groups exceeded pre test expectancy for vocabulary in twelfth grade. It should be noted, however, that the margins were narrow, 1.9 being the largest margin found.

It is the author's view that the comparatively large distances between post test expectancies and twelfth grade scores for the Middle and Low groups lend support to earlier arguments on the lack of validity of the vocabulary score as a true measure of vocabulary improvement during the reading program.

Total comprehension

The High and Middle groups both exceeded pre and post test expectancies on this measure. Both groups showed slight losses from pre to post test. The Low group, which alone showed progress between pre and post test on total comprehension, exceeded pre test expectancy; but it did not attain the higher post test expectancy.

Again, margins between pre and post test expectancies and actual twelfth grade scores were small in terms of

raw score.

Total reading

According to Table VII, page 82, all groups exceeded pre test expectancy in total reading on the test taken in twelfth grade; no group attained post test expectancy.

The difference between post test expectancy and actual twelfth grade score on this measure followed a pattern consistent with intelligence. The High group's actual score was only .3 of a point short of post test expectancy; the Middle group difference was 1.4 points; and the Low group's score was 4.2 points less than post test expectancy.

Conclusions

Since expectancy is based upon percentile rank, it is not surprising that this approach did not reveal anything really new about the value of the program studied. Again the rate score in twelfth grade seems to be positively and rather strongly associated with the progress made during the reading program. Other scores' excess over pre test expectancy can hardly be credited to the reading program on the basis of the narrow margins found. The expectancy approach would suggest that the small differences could have been caused by either lack of association between the reading program and overall reading progress or a lack of adequate range in the test administered.

CHAPTER VI

SUMMARY AND CONCLUSIONS

The author was a reading teacher at East High School during the two years following the administration of the reading program which is the subject of this study. Students who came to the author's attention later in their high school careers did not seem to measure up to the standards which would be expected on the basis of the published reports. This concern with later reading ability led to the present examination of the validity of the Summers studies and of the long-term value of the sophomore reading program reported upon in those studies.

Methods of Evaluation

Reading test scores of the 1964 senior class of Duluth East High School, Duluth, Minnesota, were studied with the aim of evaluating the reading program in which these students participated as sophomores.

Three major approaches to evaluation were undertaken--mean score growth, percentile rank growth, and expected twelfth grade skill as measured by standardized tests. In addition, a correlation study of limited scope is appended.

Results of the Study

The following findings are based upon facts presented

in this paper. However, some of the conclusions are subjective in the sense that they are not supported by statistical studies of significant difference.

1. Reading skill gains during the reading program, with the exception of comprehension in story materials, were impressive.
2. The accelerated growth in reading skill experienced in rate, vocabulary, and study-comprehension during the reading program was not continued in the two years following the close of reading instruction.
3. Students did not maintain a rate of growth sufficient to maintain the percentile rank standing they achieved in rate and vocabulary at the close of reading training.
4. The results of the analyses suggest that the program was probably most beneficial and appropriate for the students of greater intellectual ability.
5. The reading program effectively increased the reading rate of students of all levels of intellectual ability.
6. Other than in rate, the improvement in reading skill found between the beginning of grade ten and the middle of grade twelve cannot be positively

credited to the reading program.

7. The improvement in reading skills between the beginning of grade ten and the middle of grade twelve cannot be positively disassociated from the reading program on the basis of results found in this study.

8. The impressive gain in rate reported in Summers' East High Reading Center Report was accompanied by regressive factors not apparent on the surface of that report.

9. Summers' Reading Teacher article which implied that comprehension gain and rate gain occurred concommitantly in the program was in error.[1]

---

[1]In the Reading Teacher article, "Evaluation of Reading Gains in a Secondary School Reading Laboratory," Summers only reported rate, vocabulary, and comprehension. Although the article didn't indicate which measure of comprehension was meant, a comparison with the results of this study indicates that the only comprehension which could have been meant was study-type comprehension. The comprehension gain reported was for a section of the test unrelated to the rate gain shown. The present study shows that comprehension losses accompanied the very substantial rate gains.

Summers stated on page 258 of the work cited, "The total group made significant gains on all three sub sections of the standardized test." The statement is true if the influence of story comprehension upon total comprehension is ignored. However, there is the clear implication that all sub-tests were reported upon in the article. This, of course, was not actually the case. In addition, although the test is standardized, Summers seems to imply that all scores his report covers are standardized. However, there

Limitations of the Study

The value of this study has been severely limited by the lack of a control group. However, inconclusive results similar to those found in the present study are quite commonly found in the literature relating to follow-up studies of reading programs even when a control group has been used. Some such studies have been reviewed in chapter two.

Although the lack of a control group is regrettable, it was inevitable in any long-range study of the existing sophomore reading program at East High School. The program was already fully established prior to the first test reported in this study. Once established as an integral part of the curriculum, it is almost impossible to deny such a program to some students in order to establish a control group.

A question might be asked about the appropriateness and reliability of the tests used to measure students' reading changes.

As the program was described in chapter three, it does not seem unreasonable to expect the reading program to

---

is no standardization for study-type comprehension alone.

It appears that in both reports Summers' enthusiasm and optimism over the successes of the reading program led him to hastily drawn conclusions.

enhance skills in rate, vocabulary, and comprehension. About the only area mentioned by Summers which the tests apparently would not measure at all was the work aimed at improving spelling. The Diagnostic Reading Tests, Survey Section, must be considered a reasonably valid instrument for measuring that which the reading program attempted to accomplish.

The test manual for the reading tests administered lists the reliability of the tests as follows: rate .80, story comprehension .74, vocabulary .89, total comprehension .83, total reading .91.[2]

Although one might wish for higher reliability, the level of reliability found for this test is not bad as reading tests' reliabilities go. The test might be more legitimately criticized in the area of measuring achievement of high ability students.

In many areas touched by the test, the High group did not appear to make a properly proportioned gain in skill when compared with students of lesser ability. There were many students of high intellectual capability who measured at test ceiling on one or more of the sub-tests. The test,

---

[2]Committee on Diagnostic Reading Tests, Inc., Diagnostic Reading Tests, Survey Section, Directions for Administering, pp. 8-9. Published by the Committee on Diagnostic Reading Tests, Inc., Mountain Home, North Carolina, 1947.

therefore, did not adequately measure many of these brighter students and the comparison of their gains with those of lower ability was not as complimentary to the High group's progress as probably should have been the case.

Curricular Implications of the Study

The ambiguity of this study's results are such that experimentation with the East High School sophomore reading program deserves consideration. Although the present study does not in itself prove or disprove the value of the program, it does provide a point of departure for future studies of reading programs at East High School. Those taught in a manner similar to that used in 1962-1963 can act as a control group for experimentation, while the experimental group would act as a control group for students taught in the same manner as the sophomores of 1962-1963.

Should no differences be found over a period of time and a variety of experiments, temporarily dropping the reading program and only testing at the same times would appear justifiable. Under those circumstances, dropping the reading program would be the only means of determining with reasonable accuracy the value of the reading program for East High School's sophomores.

The results of this study suggest that schools view the reading center approach to senior high school reading

programs with caution. The results of establishing such a center are likely to be either the concentration of reading instruction into a small period of time or the exclusion of large numbers of students from systematic reading instruction. Exclusion of large numbers of students because they are presumed to be beyond help is now a discredited concept in secondary reading. The results of this study suggest that concentration of reading instruction into a short time period is, at the least, a questionable practice.

Suggestions for Future Research

The inconclusive results of the present study, as well as similar results in many of the studies reviewed in this paper, suggest the need for carefully controlled, long-term studies of senior high school reading programs.

The difficulty of measuring students of high and low intelligence levels on a scale which is equitable for both groups is not a problem unique to this study. However, the results which have been reported in this study provide evidence of the necessity for attempting to construct an equitable measuring device for the field of reading research at the secondary level.

Until a single equitable scale for measuring progress is available, future research in secondary reading might be

better served by using separate instruments to measure the students of high and low ability. Although progress as measured by test scores would not be directly comparable with one another, such an approach would eliminate the ambiguity of interpretation caused by bright students' hitting or nearing test ceiling and slow students' inability to finish test responses within the allotted time.

BIBLIOGRAPHY

SELECTED BIBLIOGRAPHY

Barry, Robert F., and Paul E. Smith. "An Experiment in Ninth-Grade Reading Improvement," Journal of Educational Psychology, 45:407-414, November, 1954.

Berkey, Sally. "Reading and Study Skills Program in a High School District," The Reading Teacher, 16:102-103, November, 1962.

Bish, Charles E. "An Experiment in Reading Improvement," Bulletin of the National Association of Secondary School Principals, 36:89-96, January, 1952.

Bland, Phyllis. "The Development of a Reading Program in the Evanston Township High School, Evanston, Illinois," in Helping High School Students Read Better. Chicago: Science Research Associates, 1954, 118-121.

_____. "A High School Developmental Reading Program," The Reading Teacher, 8:146-152, February, 1955.

Blough, Richard. "The Developmental Reading Laboratory--A Training and Proving Ground for Readers," Bulletin of the National Association of Secondary School Principals, 40:38-41, November, 1956.

Committee on Diagnostic Reading Tests, Inc. Diagnostic Reading Tests, Survey Section, Directions for Administering. Mountain Home, North Carolina: The Committee on Diagnostic Reading Tests, 1947.

_____. Diagnostic Reading Tests, Survey Section: Upper Level. Mountain Home, North Carolina: The Committee on Diagnostic Reading Tests, 1947.

Dobrin, Ruth M. "The Massapequa Story: a Pre-College Reading Program," Journal of Developmental Reading, 4:159-172, Spring, 1961.

Dehl, Kermit. "Reading Improvement Practices in the Oak Park and River Forest High School, Oak Park, Illinois," 118-121 in Helping High School Students Read Better. Chicago: Science Research Associates, 1954.

Engelhardt, Reata M. "Speed Is Not a Naughty Word," Journal of Reading, 8:330-331, April, 1965.

Freed, Ruth. "Developing a High School Reading Laboratory," Journal of Secondary Education, 40:303-309, November, 1965.

Glock, Marvin, and Jason Millman. "Evaluation of a Study Skills Program for Above Average High School Pupils," Journal of Developmental Reading, 7:283-289, Summer, 1964.

Gowan, J. C., and R. W. Scheibel. "The Improvement of Reading in Gifted Children," Educational Administration and Supervision, 46:35-39, January, 1960.

Jackson, Jeanne R., and Henry Dizney. "Intensive Vocabulary Training," Journal of Developmental Reading, 6:221-229, Summer, 1963.

Kendall, Rose. "Give Them a Chance," Journal of Reading, 8:326-329, April, 1965.

Leamnson, George F. "Indianapolis Produces Better Readers," School Executive, 74:64-67, December, 1954.

Marquis, Bettylee F. "Developmental Reading--New Albany High School," Journal of Developmental Reading, 7:58-62, Autumn, 1963.

Melton, Howard B. "Reading Retention of Eleventh-Grade Students," Journal of Developmental Reading, 2:26-33, Summer, 1959.

Perry, Harold J. "The Developmental Reading Program in the Highland Park High School, Highland Park, Illinois," 122-127 in Helping High School Students Read Better. Chicago: Science Research Associates, 1954.

Simpson, Elizabeth A. Helping High School Students Read Better. Chicago: Science Research Associates, 1954.

Summers, Edward G. East High Reading Center Report. Duluth, Minnesota: Duluth Public Schools, 1962.

Summers, Edward G. "Evaluation of Gains in a Secondary School Reading Laboratory," The Reading Teacher, 17:255-259, January, 1964.

Thornton, Cecil M. "Two High School Reading Improvement Programs," Journal of Developmental Reading, 3:115-122, Winter, 1960.

Tormey, Mary K., and Walter G. Patterson. "Developmental Reading and Student Evaluation," Journal of Developmental Reading, 2:30-43, Winter, 1959.

APPENDIX

CORRELATION STUDIES

As noted in chapter five, the correlation studies were not among the original plans for evaluation for this paper. As will be seen below, the validity of the results found is such that inclusion within the body of this report did not seem proper.

Method of Finding Correlations

A study was made of correlation between pre to post test changes in scores and pre to follow-up test changes in scores; and a study was made of correlation between pre to post test changes in scores and post to follow-up test changes in scores. The Pearson product-moment was the means of correlation used. The preliminary computations were taken to two decimal places for all scores except rate and rate of comprehension. For these two measures computations were only taken to the nearest whole number.

The identical method was used in selecting the ten students' scores found in the random sample of each group. Summaries of students' scores were on individual sheets of paper. The sheets within each group were not in any pre-arranged order. For the High group every seventh paper was pulled from the pile; for the Low group every ninth paper was pulled from the pile; for the Middle group every

eighteenth paper was pulled from the pile. The results of this method were random samples of ten sets of scores for each of the groups.

Under the circumstances, sampling was the only practical means of doing a correlation study. The calculations are too time consuming for a study of all the available scores to have been a practical undertaking. The samples were deliberately kept small to facilitate handling the computations involved.

As there is no statistical proof that the samples drawn are identical to the groups they represent in characteristics of score changes, the correlation results should be interpreted cautiously. The variations between total group and sample group means at the various testing points are presented in Table X at the end of this appendix.

Results of Correlation Studies

Correlations of less than .75 will not be discussed. Since the correlation study here is based upon identical students taking identical tests, any correlation of less than .75 has to be considered relatively useless in evaluating the reading program.

Two studies of relationship were undertaken. The first correlation discussed here was for pre to post changes with pre to follow-up changes. The second cor-

relation compares program changes with post-program changes.

Where the first correlation study may be an indicator of the appropriateness of the program to students' needs, the second correlation is a more likely indicator of the follow-through value of the program.

Pre to post-pre to follow-up correlations of change

It must be remembered that in this correlation a part is being compared to a whole.

Table VIII, page 101, shows the correlations found for all groups. Six of the eighteen possibilities produced a correlation exceeding .75. All correlations were positive.

The vocabulary correlation was relatively high for all groups. It may be remembered that no group showed regression at any point on this measure. It may also be remembered that the pre to post gain was relatively large on this measure. Only the High group had high correlations on other score changes.

The fact that high correlations were found for the High group on rate, rate of comprehension, and total reading would seem to suggest that the program was most appropriate and helpful for the High group. This correlation study suggests that the program's effect upon the Middle and Low groups is open to serious question.

TABLE VIII
PRE TO POST-PRE TO FOLLOW-UP
CORRELATION OF CHANGES FOUND

Skill	High group (n=10)	Middle group (n=10)	Low group (n=10)
Rate	∓ .88	+ .12	+ .05
Rate of Comp.	+ .76	+ .10	+ .40
Story Comp.	+ .02	+ .42	+ .46
Vocabulary	+ .80	+ .80	+ .87
Total Comp.	+ .05	+ .22	+ .74
Total Rding.	+ .77	+ .73	+ .55

Pre to post-post to follow-up correlations

The correlations found between changes occurring during the reading program and changes occurring in the following two years are shown in Table IX, page 103. A glance at that table shows that many of the correlations for skill changes during these intervals were negative, and several of the positive correlations were found where there was loss during the reading program.

Only two of the eighteen correlations exceeded .75 and both were positive. The Middle group pre to post test total comprehension change (score loss) was highly and positively associated with post program gain. The Low group pre to post test change in rate (gain) was highly and positively associated with post program loss.

Clearly, this correlation study does not support the conclusion that the reading program had residual value in the sense that students' later growth in skills was closely related to growth during the program.

Conclusions

Both correlation studies add to the trend of evidence found in mean and percentile rank comparisons which suggests that the reading program may have been most appropriate and beneficial for the High group, while less useful for the Middle and Low groups.

TABLE IX
PRE TO POST-POST TO FOLLOW-UP
CORRELATION OF CHANGES FOUND

Skill	High group (n=10)	Middle group (n=10)	Low group (n=10)
Rate	+.61	+.43	+.93
Rate of Comp.	-.24	+.64	+.69
Story Comp.	+.51	+.69	-.12
Vocabulary	+.22	-.34	-.02
Total Comp.	-.23	+.79	-.06
Total Rding.	-.01	-.29	+.53

The correlation studies seem to indicate that tenth to twelfth grade gains may well be attributable to factors other than the reading program.

TABLE X
THE AMOUNT AND DIRECTION OF THE VARIATION OF THE MEANS FOR THE SAMPLE GROUPS FROM THE MEANS OF THE TOTAL GROUPS WHICH THEY REPRESENT

Skill	High group	Middle group	Low group
RATE			
Pre	-17	+1.7	+13.6
Post	+7.9	-8.4	+41.3
Follow-up	+28.1	-9.7	+24.0
RATE OF COMPREHENSION			
Pre	-11.5	-13.7	+9.0
Post	+31.1	-22.0	+25.0
Follow-up	+43.8	-25.4	+11.3
STORY COMPREHENSION			
Pre	+.2	-1.0	+.2
Post	+.5	-.8	+.2
Follow-up	+.4	-1.0	-.2
TOTAL COMPREHENSION			
Pre	0.0	-.4	-.3
Post	+.7	-1.6	-.3
Follow-up	+.7	-2.1	-.4
VOCABULARY			
Pre	-1.2	-1.2	+.4
Post	-.8	+1.1	+.2
Follow-up	+.1	+2.3	0.0
TOTAL READING			
Pre	-1.1	-1.6	+.1
Post	-.2	-.5	-.1
Follow-up	+.8	+.2	-.4