5-1962

An Analysis of the Achievement Gains Made By Students in Ability-Grouped Vs. Random-Grouped Classroom Units

Luan H. Ferrin
Utah State University

AN ANALYSIS OF THE ACHIEVEMENT GAINS MADE BY STUDENTS IN ABILITY-GROUPED VS RANDOM-GROUPED CLASSROOM UNITS

by

Luan H. Ferrin

A thesis submitted in partial fulfillment
of the requirements for the degree

of

MASTER OF SCIENCE

in

Educational Administration

Approved:

UTAH STATE UNIVERSITY
Logan, Utah

1962

ACKNOWLEDGMENTS

Writing a thesis is a task which requires the energies and attention of many people. I express my deepest appreciation and thanks to the following people:

Doctor Walter Borg, chairman of my committee, who conceived and directed the larger study of which this treatise is but a small part, for his patience and help over the period of time required for the completion of this project.

Doctor Basil Hansen and Mr. Arthur Jackson, the other two members of my committee, for their encouragement and direction.

To my wonderful wife, Wilma, for her help with the typing and editing and for the constant encouragement she provided.

To Mike and Ann for their patience with their daddy while the work was being accomplished.

To my many friends and associates for their suggestions and encouragement, and to all others who have helped in any way to bring this thesis to completion.

Luan H. Ferrin

TABLE OF CONTENTS

LIST OF TABLES

CHAPTER I

INTRODUCTION

Because of the rapid advancements being made in the field of knowledge, educators, as well as those in other fields, must periodically take inventory. Current practices, policies, and methods must be carefully scrutinized to determine if they are the most effective. The group or class method of instruction is one such area.

The current philosophy of education held in many parts of the United States today places a great deal of importance upon the individual child. Numerous programs have been inaugurated to give the individual child as much attention as possible and still be able to have a class large enough to be practical financially. This task becomes increasingly difficult when the range of abilities within each classroom is so great. It isn't uncommon in the upper elementary and secondary classes to find a spread of from six to nine years difference in ability or achievement within one classroom.

Not only do we have the problem of range within the classroom, but with the increasing school population of today, classes have grown to a prohibitive size. Add these and other problems that stem from the pressures of present day society together, and even with the best possible teacher, we get only average results.

Many experiments and programs have been dedicated to the solution of this problem.

Statement of Thesis Problem

Origin and nature of the problem

The ability grouping system has for years held out a possible solution to these problems by providing a system that recognizes and is adaptable to individual differences in an economically feasible manner. In spite of the promises held out by ability grouping, relatively few schools are currently using this technique. This is partially because the research in the field has not demonstrated that ability grouping does the things claimed for it by its proponents. There are many phases of the question where research is completely lacking.

This study germinated when the administration of a local school district (hereafter referred to as district "A") embarked on a program of ability grouping.

On the basis of the composite scores of the California Achievement Test Battery, Form W, plus the evaluation of the teacher and principal, the students at each grade level were grouped into three categories: developmental, regular, and accelerated. They were then placed in a classroom with others of the same classification. Material in the curriculum was then adjusted for each group.

Adjacent to district "A" is a district similar in geographic location (hereafter referred to as district "R")

where the program of grouping according to age and grade level was being maintained. District "R" is a city district and district "A" is a county district, but as the suburbs of the city fall within the county boundaries, it was possible to match the schools in a reasonably accurate manner. The schools in the study were selected on the basis of recommendations made from both districts. Later, the Sequential Test of Educational Performance was administered and the schools were found to be comparable.

Many people, however, feel that the strengths of the program do not justify its use because the weaknesses overshadow the good that can come from it. Wallin (1956) reports some of the arguments for and against ability grouping:

For:

1. We have no right to neglect the dull.
2. More individual attention can be given.
3. Children can learn at their own rate.
4. Students learn more thoroughly if they do it at their own rate and they also adjust better.
5. Children are happier with children of their own level and learn more from them.
6. They acquire more confidence in their own ability.
7. A goal within reach serves as an incentive to them.
8. All children experience success.
9. They can contribute to discussions.
10. Children become discouraged by unequal competition with the bright and often become behavior problems.
11. The bright accomplish more in grouped classrooms.
12. They are not bored or discouraged by needless repetition.
13. They have fewer opportunities to show off.
14. Teachers can make better adjustments to individual differences.

Against:

1. Children learn from each other at all levels of ability.
2. The dull derive much of social and emotional value from the bright.
3. The dull secure stimulation and help from the bright.
4. Children of different abilities need to learn to work together in school as they face this in life situations.
5. Children need a normal balanced situation.
6. Sectioning stigmatizes the dull.
7. It places too much emphasis on mental ability.
8. Grouping can never be homogeneous.
9. The same curriculum is often used for both groups.
10. The teachers for both groups often lack the training and materials or sympathy for the problems of the fast or the slow.

Many of these objections to and contentions for sound very valid and bear careful scrutiny.

Because of the conflicts of opinion and because there are many areas where research is completely lacking, the ability grouping program is not in wide use. One such area where research is lacking is in terms of the internal scores of the evaluating instruments themselves. Nowhere is the literature surveyed was a study found that even claimed to measure these differences.

This study was set up for that purpose. It is designed to measure the advantages and disadvantages of the ability grouping program as it affects certain types of students. It focuses attention on those students whose scores on the California Achievement Test Battery are homogeneous or heterogeneous. That is, their scores on the three sections of the test battery are relatively close or are widely spread. The possibility of measuring other variables

became evident as the planning progressed. Sex differences were considered as were interdistrict and intertype differences.

Using district "A" as the experimental district and the adjoining district "R" as the control district for the experiment, the possibility for valid research and a possible answer to many of the problems of grouping was excellent.

Under the direction of Dr. Walter R. Borg, head of the Bureau of Educational Research at Utah State University, the study has taken form and at the present time is progressing according to schedule. This portion of the major research concerns itself with the first year of the study. The entire program will last for three or more years.

Hypotheses

I. To measure and compare sex differences in the two systems. These comparisons will deal with the following hypotheses:

A. Boys are significantly more homogeneous in terms of their internal achievement scores on the California Achievement Test Battery than are girls.

B. Boys at the various levels of homogeneity do not gain significantly more than girls at the same level.

C. There are not significantly more homogeneous boys in district "A" than in district "R."

II. To measure and compare gains made by pupils in the two systems who fall into the various internal homogeneity classifications. These comparisons will deal with

the following hypotheses:

A. Students whose internal homogeneity scores on the California Achievement Test Battery are homogeneous do not achieve significantly more than those that are regular or heterogeneous.

B. Students whose internal homogeneity scores on the California Achievement Test Battery are regular do not achieve significantly more than those that are homogeneous or heterogeneous.

C. Students whose internal homogeneity scores on the California Achievement Test Battery are heterogeneous do not achieve significantly more than those that are homogeneous or regular.

III. To measure and compare gains made by pupils in the two systems. These comparisons will deal with the following hypotheses:

A. Students who are homogeneous in terms of their California Achievement Test Battery scores in the ability-grouped classes achieve significantly more than do homogeneous students in a random-grouped situation.

B. Students who are "regular" (neither homogeneous or heterogeneous as determined by the scale) do not achieve significantly more in the ability-grouped situation than those in the random-grouped situation.

C. Students whose initial California Achievement Test Battery scores were heterogeneous achieve significantly more in a random-grouped situation than do the heterogeneous students in an ability-grouped class.

Definition of Terms

"Ability grouping" is the name given to the system in the study that channels a student into a classroom on the basis of achievement and ability plus evaluation by teacher and principal along with other students of similar classification.

"Homogeneous" as defined in this paper means alike or similar. More specifically, those students whose test scores on the California Achievement Test Battery in the areas of arithmetic, language, and reading were very close together or alike were classified as homogeneous.

"Regular" students were those whose scores on the California Achievement Test Battery were neither closely gathered nor widely separated but fell between these two classifications. On the total distribution of scores, they represent approximately + or - one sigma from the mean.

"Heterogeneous" as used herein signifies different or varied. Students whose three scores on the California Achievement Test Battery were widely separated were classified as the heterogeneous sample.

"Random-grouped" classes are those classes that are grouped according to age with little attempt being made to structure it beyond that medium. Each child in the population has an equal chance to be chosen.

"Internal homogeneity scores" are those scores based on the difference between the highest grade placement score

earned by a pupil and the lowest grade placement score earned by the same pupil on the three different sections of the California Achievement Test Battery.

"Analysis of covariance" is the statistical tool used to equate two groups that are initially unlike. This method allows for correlation between initial and final scores and makes possible the adjustments in final or terminal scores which will allow for differences in some initial variable.

CHAPTER II

REVIEW OF LITERATURE

The literature written about ability grouping and related subjects is voluminous. A review of the results of these studies, however, leads one to believe that there certainly is a need for strong, objective, conclusive type studies that make it possible to state definite conclusions. As Wyndham (1934) says: "The first general impression that one gains from these studies is that . . . they raise more issues than they settle."

Many people express themselves strongly on the subject with very little objective data to substantiate their feelings. The interest manifest by these studies is indicative of the need for work in this area.

Ability grouping in various forms has always been a part of our educational system. The class or chronological age-group method, itself, was an attempt to place students in a general classification where they could be taught similar material. Elective classes in the secondary schools group students. Some students elect foreign languages and higher mathematics. This brings together students of greater ability. The students who can't make these courses are given substitute offerings. Still further, students

elect music, art, and other special fields which give them security. This again groups students.

Even though grouping is not new, when an attempt is made to change methods of grouping, discussion usually follows. Let's turn our attention to some of the studies that have been made to pick up some of the reasons for and against the use of ability grouping.

For and Against Ability Grouping

McGaughy (1930), as did Wallin (1956), lists some of the objections to ability grouping and some of the reasons proponents push for it. They are as follows:

For ability grouping

1. Because the students are similar in achievement in school studies, they are easier to teach.
2. If the group is relatively alike, the curriculum can more easily be adapted to that group.
3. With similar rates of speed, students are happier together and enjoy school life more.
4. The percentage of failure is lower in homogeneous groups.

Against ability grouping

1. Because "school is life and not just preparation for it," ability grouping creates an unnatural condition.
2. Children in slower groups are marked as "dullards" or "dumbells."
3. If teaching special groups is so important, a person should be specially trained for his or her level. Yet, because most teachers don't like to teach the slow class, a rotation from year to year takes place.
4. Often the brightest students develop a "superior attitude."

These are just a few of the pros and cons of the subject. Let's go further. An opinion poll (1955-56) was made by the

Nation's School. They ask the question, "Should children be grouped through the early years on the basis of ability rather than according to the typical age-grade system?"

A nationwide sampling of school administrators was used.

The administrators indicated their feelings by a 60 per cent vote against and a 40 per cent vote for the program. Those who favor ability grouping point out that it is more likely that an outstanding pupil will get the attention he needs. They also state that it probably works better where there is superior teaching and in larger cities where generally there is less parental friction than in small towns. Many of those who favor the age-grade grouping have no objection to grouping within the classroom. They suggest that ability grouping is wrong because pupils learn much from dealing with others of contrasting ability.

Alice Keliher (1931), in her work on ability grouping, lists five assumptions which the grouping implies. These will now be considered with studies that give the opposing views, as given by Hammond (1959).

> 1. Intelligence is so adequately measured by verbal intelligence tests that the results may serve as basis for action which concerns the whole individual. The contention that types of intelligence needed in school are adequately measured by tests has probably arisen from definitions of intelligence as ability to learn, interpreted narrowly as acquisition of academic skill. This position also assumes that intelligence functions consistently. The evidence refutes this statement. The Terman Group Test of Mental Ability includes ten categories such as information, word meaning, logical selection, arithmetic, analogies, etc. The IQ and MA are

both derived from averages of different mental functions, yet individuals who achieve the same composite of average results do it by many combinations of specific abilities.

There are those who believe that the educational age forms a sound basis for classifying individuals. It must be remembered, however, that education includes more than the enumerated components of the educational age. Reading age is important, but so are reading attitudes and dispositions. The restriction to limited academic attainments neglects many physical, social, and emotional traits which make up the whole individual. On this point Keliher (1931) says: "It is simply a statement of fact that the combined measures of verbal intelligence and the academic skills, plus a vague factor of Teacher's Judgment, which may or may not concern itself with other than academic skills, do not represent more than a small portion of the traits characteristic of an individual. For this reason, the use of these bases for any action which concerns the whole individual, when traits other than those measured are to be affected, is without justification."

2. A further assumption is that homogeneity of grouping reduces the range of variations with a grade. In regard to the reduction of variations, Burr (1931) found that after grouping had been carried out, four-fifths of the total range of ability in the original undivided group remained in each of the so-called homogeneous groups. In separate cities the overlapping ranged from 68 per cent of the total grade range.

In an earlier study Courtis (1923) said: "Sectioning on the basis of intelligence is a device for securing homogeneous groups; yet measurement of the achievements and growths of individuals in sections of supposedly equal intelligence proves that not all the bright children succeed and that not all the dull children fail. There is both success and failure in each group to such an extent that in the highest and lowest fifth of 4,000 first grade children the number of individuals having identically the same scores in a reading test at the end of the semester were recently found to be one-half the total number. Further, there are some data which suggest that for any large group of children the total distribution and the median scores are the same whether the individuals are taught in undifferentiated sections or in

classes carefully sectioned on the basis of intelligence.

The conclusions being drawn from these data are two: first, that intelligence is but one of many factors affecting a child's success, and second, that individual differences are so great that no method of work can be made effective which does not provide for the complete adjustment of assigned tasks to the nature and powers of each child each day. Sectioning on the basis of intelligence scores is apparently proving to be a temporary expedient, a more refined method of grading, but not an ultimate solution."

3. Perhaps, the most important assumption is that homogeneity of grouping tends to bring superior learning results. Cornell (1936) reports that, "a review of objective results of ability grouping leaves one convinced that we have not yet attained unequivocal experimental results that are capable of wide generalizations." Wyndham (1934) says, "the first general impression one gains from these studies is that . . . they raise more issues than they settle."

Miller and Otto (1930) analyzed thirteen experimental studies of homogeneous grouping, and conclude their summary by saying: "If one were to make a final summary statement about the studies represented . . . one would have to say that, so far as achievement is concerned, there is no clear-cut evidence that homogeneous grouping is either advantageous or disadvantageous."

Keliher (1931) comments on the expectations of teachers regarding these groups: "The degree of expectedness of improved achievement or poor achievement is a matter of concern here. It may be possible that the teaching attitude of expectedness of results is keyes to the supposed mediocrity of each intelligence level . . . Certainly the attitude of the teacher concerning what she may expect from her class is a most potent factor in the attainment of results. The teacher who is complacent with regard to the limitations of her slow group will not put forth the effort or show the interest required to elicit the highest possible performance from these children. Therefore, it is probably true that equalizing this factor, or accounting for it in testing out the results, would in many cases actually change the results."

There are those today who propose this grouping in order to care for the gifted. One of the most consistent results has been the

possibility of increased speed in covering a given amount of work on the part of bright children. But is this adequate provision for the gifted?

4. Another important assumption is that homogeneity of grouping tends to make superior provision for individual differences. This type of grouping can easily lead a teacher to be less alert to detect and provide for individual differences. The use of an average or averages as the basis for grouping and the concern for uniform achievement tends to turn the thinking of the teacher away from the individual toward average results.

Alberty points out: "Fundamental to any program based upon ability grouping is the assumption that learning takes place more effectively if the range of differences in pupil ability is materially reduced, so that learning activities that will be appropriate for the group as a whole may be selected . . . Yet the fact remains that the device itself lends itself to the facility of uniformity of assignment and instruction. The aspects of such mass instruction will be less obvious when pupils are grouped more homogeneously. Consequently, the teacher will be less likely to recognize and provide for individual differences."

5. There are those who say that homogeneous grouping offers more chance for success and happiness, eliminates snobbishness and conceit of bright pupils, and that slow children do not experience the discouragement of daily failure. These ideas are based on the assumption that Homogeneous grouping provides for better attitudes in pupils. On this point, a study was reported by Keliher (1931) in which she observed the response of children in one sixth grade and two eighth grades, grouped heterogeneously. The results showed the tendency for the brighter children to remain in the upper 75 per cent of responses. The important point in relation to suppression of children of low intelligence, however, is that for two eighth grades, the children of the lowest 30 per cent in intelligence are as likely to be in the upper 30 per cent in responses as they are to fall in the lower one-half. In the three classrooms observed in which progressive practices were followed, discouragement and suppression do not necessarily occur in mixed groups in any fixed relation to intelligence.

These studies emphasize the mixed feelings held by educators on the subject.

Bases for Grouping

Certainly finding a basis for grouping presents a problem to all who would attempt to group. Kefauver (1929) reports on a sampling of seventh grade students who were grouped in Fresno, California. The bases used were: (1) average of their school marks received in the fifth and sixth grades, (2) teacher's estimate of the students application to study, (3) teacher's estimate of capacity, (4) Multi-mental Test scores, (5) intelligence quotient, (6) Thorndike-McCall Reading Test T scores, (7) Woody-McCall Arithmetic Test scores, (8) Monroe Reasoning Test scores.

It was found that the most significant single source of information for predicting success in the first year of junior high school is the judgment of the teachers in the elementary schools. The general intelligence test is the most accurate of the tests for predicting general success, but it is superceded by special achievement tests for predicting success in individual subjects. The general achievement test covering the content of a number of subjects shows a high relationship to general success.

Another study by Washburn (1924) reports the attempt to determine gifted children and group them by National Intelligence Tests. After the tests were administered, the highest one-fourth were selected. Their I. Q.'s ranged from 123 to 166. There were 192 students within this range. The lowest quartile was grouped together as was the center 50 per cent.

They concluded that: (1) Gifted children allowed to move through school at their own rate make, as a whole, distinctly better progress than children of lower I. Q.'s. (2) There is a wide range in the rate of progress of gifted children—so wide that the lower half of the group actually progress less rapidly than an equal number of children from the top of the middle group. (3) Gifted children not only differ in average rate of progress but in the subjects in which they excel. (4) Any attempt to group children by intelligence quotients will result in the misplacement of nearly half of the gifted group and often an equal number of the middle group.

Both of these studies emphasize the problems involved in setting up valid criteria for grouping. Some studies, however, have a brighter picture. Roberts (1947) reports on a study where the students were placed by the principal into groups according to their achievement in reading and arithmetic. The gains made were small and the study had many variables which could have been equally as important to the study as the grouping method used.

In 1920 a study was made using the Illinois Intelligence Test reported by Theisen (1922). The students were grouped on the basis of this test giving some weight to their previous records. The tests were administered in June after six months of the program were complete. The sections that made the higher intelligence test scores in each school excelled in scholarship. Intelligence and achievement were

correlated and found positive. The highest correlation was between intelligence and arithmetic. Reading ranked next with language third.

These studies seem to justify the assumption that intelligence and achievement are valid bases for grouping students.

One other report by Kozal (1958), using the opinion expressed in a discussion group, lists four methods of grouping in their preferred order: (1) group I. Q. tests, (2) low and high standing students in individual groups, (3) reading ability, (4) emotion stability, industry, and study habits.

While these opinions are not documented by experimental studies, they do reflect the thinking of some top rated administrators. Some of the most documented work in this area has been done by Turney (1931). He quotes studies by Rankin (1931) and Billett (1932) and comes up with a list of variables in grouping.

These variables have been grouped in seven categories for convenience, and certainly the list doesn't claim to be exhaustive.

A. Physical Development
   1. Chronological Age
   2. Physical Maturity
   3. Physiological Maturity
   4. Health
   5. Height
   6. Weight
   7. Anatomical Age

B. Intelligence
   8. Intelligence Test Results
      a. Raw Score
      b. Mental Age

c. I. Q.
9. Teachers' Ratings, Singly or Average
a. Of Ability to Learn
b. Of Section to Which Pupil Belongs
10. Probable Learning Rate

C. Achievement
11. Achievement Test Results
a. Educational Age
b. Achievement Quotient
c. Subject Age or Subject Quotient
d. Raw Scores on One or More Subject-Matter Tests
12. Teachers' Marks in One or More Subjects
13. Rank in Class

D. Motivation
14. Ratings or Judgments on Traits (like Industry and Application)
15. Achievement Quotients or Similar Indexes (see 11 above)
16. Rank in Class (see 13 above)

E. Social Factors
17. Social Age or Maturity
18. Home Environment

F. Special Abilities and Interests
19. Prognostic or Placement Test Results
20. Special Ability Tests (as in Music)

G. Special Disabilities
21. Defective Vision or Hearing
22. Physical Deformity
23. Speech Defects

These studies again emphasize the complexity of the problem.

Ability Grouping and Achievement

We have now considered the bases for grouping. Let's weigh the literature to see if ability grouping actually aids academic progress, for in this area proponents make their greatest claims.

A study reported by Riley (1956) was revealing. There were 154 sixth grade students who were grouped using general achievement, standard tests, teachers opinion, reading ability, creative ability, and I. Q. They were divided into

four groups. Standard tests were given again near the end of the year. The results showed 14 to 54 months progress made during a nine month period. The class median was 17 months higher than the median for the first test. The slower classes were able to achieve and were, therefore, happy. Other results reported include: Sports programs were carried on without problems, all levels winning some games. The social development of the slow groups was significant. It was also emphasized by this study that teacher and principal attitudes and qualifications are very important.

A project with a weak design but nevertheless applicable to our concern here was reported by Nash (1942). Each year at a business school small groups representing the highest and lowest levels of student ability were formed. One of these, a low class of 15 members, was reported in this summary. A special core curriculum was set up and oral work was used because of the poor reading ability of the group. Typing was taught as part of the modified curriculum. This group was held intact for three years. Of the 15 original members, 11 fulfilled all diploma requirements and graduated with their class; two left to be married; two others were placed by the school in local jobs.

It was felt by the person reporting that because of this special programing that all of the girls were met on their level. Grouping saved them from failure. With as many uncontrolled variables as were apparently evident in this study, not much weight can be given to the outcome.

One of the better studies on ability grouping was done by Barthelmess and Boyer (1932). In this study the students from five schools in an area of Philadelphia were sectioned into groups according to some measure of intelligence and achievement not mentioned in the report. In the primary grades, students were grouped according to individual examinations given by specialists. In the intermediate grades, they used highly verbal group tests. A control group was formed for each level. Five tests besides the Otis Classification Test were administered at the beginning and at the end of the year. They are as follows: Philadelphia Test of Problems in Arithmetic, Philadelphia English Test, Philadelphia Test in Fundamentals of Arithmetic, Philadelphia Geography Reading Test, and Stanford Test in Paragraph Reading.

The total 565 experimental pupils made an improvement of 12.8 months and the control group made 10.4 or a difference of 2.4 months for the first year. During the second year the 297 pupils in the experimental group improved 13.5 and the control group made 11.3, an average of 2.2 months difference.

This study attempted to control all variables. Even teachers were tested and matched as nearly as possible.

The results point out that in arithmetic, reading, and technical English skills, there is a strong statistical significant difference in favor of ability-grouped pupils.

A similar result was found by Kvaraceus and Wiles (1938) in a study conducted in a Massachusetts School District. On the bases of the Metropolitan Achievement Tests and the judgment of teachers, the sample was classified into groups X, Y, and Z according to their achievement and apparent skills in reading, English, and arithmetic. Each class spent some time with the entire group. At the end of the school year, the testing program was again administered.

In the autumn of 1937, the 75 experimental students were the lowest in the district. On retesting at the close of the 1937-38 school year, these classes in the experimental school rose from lowest in the district to fifth from the bottom.

This data indicates that more than the average pupil growth was experienced by these students in the course of that year.

Work done in the Detroit Public Schools was among the first research in the field. Vreeland (1932) and Rankin, Anderson, and Bergman (1936) have reported on the progress of the study. The study began in 1920 when all students entering the first grade were placed into X, Y, or Z groups on the basis of a group intelligence test. The superior 20 per cent were organized into X groups. The middle 60 per cent were organized into Y groups. The lower 20 per cent were grouped into Z groups. Differential curriculums were provided as needed for each group to meet interest and ability. Three plans were used, the Vertical Plan of

Grouping, the Mass Instruction Plan, and the Typical Detriot Plan. The Vertical Grouping Plan (1929) was designed to carry the principle of ability grouping farther than is true in Typical Detroit Schools. In the Mass Instruction Plan (1929) all students were taught as nearly alike as possible. Materials, methods, and standards were all the same. The Detroit Plan, or X, Y, Z plan, is outlined above.

The overall purpose of the study was to evaluate the effectiveness of the three levels of adaption to differences in bright, average, and dull pupils. The Vertical Plan illustrates considerable adjustment, the Detroit, moderate adjustment, and the Mass Plan, little or no adjustment.

The test results given during the experiment indicate a superiority for the Vertical Plan of about 20 per cent over Detroit and even more than that over the Mass Instruction. In arithmetic the Detroit was superior, about 14 per cent over the Vertical. Data was presented for 500 pupils in grades three to six.

Each of these plans (Detroit and Vertical) that use ability grouping seems to produce more satisfactory results than does the Mass Instruction Plan.

Another study in achievement that produced positive results was conducted by Hartill (1936). In grades five and six in New York City Schools the Stanford Achievement Test was given to all students in December 1931. They were then sectioned into 1, 2, and 3 groups. Then the

homogeneous group was given a differentiated course of study from February to June 1933. The heterogeneous group had the regular course of study. From September 1932 to January 1933 the groups were reversed. At the close of the experiment all students were given a different form of the Stanford Achievement Test.

They reached the conclusions as stated: Homogeneous grouping of the type arranged is better than heterogeneous grouping. Under homogeneous grouping the subject matter gains for the whole group were as large as those under heterogeneous grouping. Students under the homogeneous program received enrichment that they didn't receive under the other plan. Significant gains in the fundamental subjects were also noted by those in some of the homogeneous groups. They felt that grouping should always be flexible and temporary and that even though homogeneous groups show definite advantages, there is also a place for heterogeneous groups.

Taking the major studies as a whole, ability grouping can't be condemmed from the achievement aspects.

Problems of Grouping

Opponents of ability grouping usually don't attack it from the basis of achievement. They feel that even though the achievement is positive that other complications make it dangerous. Let's examine some of the studies that point up a few of the problems encountered.

In an interview study of 190 fourth, fifth, and sixth grade children by Luchins and Luchins (1948), the attempt was made to determine children's attitudes toward homogeneous grouping. The choice was almost unanimous in the brightest group. It decreased as it went to the average group and then to the dull group. In the dull group, three-fourths of the children prefer number one class status.

This study indicated that dull pupils appeared to feel inferior and ostracized. There was a decided stigma attached to the number two class label and strong pressure to be in the number one class. Along with these other problems mentioned, a snobbish and superior attitude was present in the number one class.

Mann's study (1957) emphasizes what Luchins calls the cast system. He attempted to find out how much carry over there was in friendship after groups had been separated by ability grouping. The procedures developed were designed to measure the social position the gifted children held among gifted as well as typical classmates. The procedures consisted of two sociometrics and a parent questionnaire. The first sociometric asked three acceptance-oriented and three rejection-oriented questions. The children were told that they might choose from any of the pupil population, kindergarten to sixth grade, attending school. To obtain rejection-oriented responses, the questions substituted the words, "least like," for the word, "like." The second sociometric was designed primarily to examine the likelihood of a gifted child choosing

a typical child in those classes which both attended. Finally, a questionnaire to be sent to parents of the workshop children was developed. Two things were to be examined through the parent questionnaire. The first was the consistency of social status a gifted child attained in and out of school. Was the most popular gifted child in school, the most popular out of school? The second thought to be examined was the belief that Pregler held concerning admissions to the Colfax School. Ordinarily gifted children in many communities are transported to a special class from various parts of the city. At Colfax, however, only those children residing within the school district which Colfax normally serves are accepted for admission. Pregler feels that such a policy would tend to develop and reinforce further the friendships that gifted children made.

The first sociometric was given to children drawn from the fourth, fifth, and sixth grades—in all, 281 children. Of this number, 67 were gifted children. These 67 came from two workshops at Colfax—the intermediate and the senior workshop groups. The intermediate workshop group consisted of 31 gifted children drawn from fourth and the lower half of the fifth grades. The senior workshop group consisted of 36 gifted children drawn from the upper fifth and sixth grades. An analysis of the results gave strong evidence that while gifted children did have visible social and academic contacts with typical children, this contact was far from real. Here gifted children, as members of the

intermediate workshop, chose other gifted children 181 times more than typical children. In the senior workshop they chose other gifted children 124 times more than typical children. Typical children, too, when they chose friends, seemed to prefer their own. Typical children from the intermediate regular classes chose other typical children 524 times more than gifted children. In the senior regular classes they chose other typical children 806 times more than gifted children. In all instances, gifted and typical children significantly chose and rejected more of their own group. The results of the second sociometric which was given to the 67 workshop children tended to reinforce the findings on the first sociometric. In the intermediate workshop, gifted children preferred other gifted children to criticize their work in music and art 71 per cent of the time; in the senior workshop they preferred gifted children to criticize this work 65 per cent of the time. The final procedure, the parent questionnaire was sent to the homes of the 67 workshop children. Parents were asked to fill in the questionnaire without consulting their youngsters. A 93 per cent return revealed that there was a substantial relationship between the friends the workshop children had in school and those they had in the community. When the acceptance choices of workshop children on the first sociometric were compared with the children listed by the parents as their child's most chosen associate, in each of the three situations, a

correlation of + .42 was found for intermediate workshop children and + .39 for senior workshop children.

They concluded that the sociometrics indicated: (1) As a group, the workshop children tended to accept and reject more workshop children than typical children. (2) As a group, typical children tended to accept and reject more typical children than workshop children. (3) In both cases there were a significant difference in the acceptance-rejection scores obtained by workshop children from typical children with whom they shared a common home room and those obtained from gifted children with whom they shared a workshop. The parent questionnaire indicated: (1) There was a substantial relationship between the friends the workshop children had in school and those they had in the community. The higher the school acceptance score the more frequent the mention of the child's name on the parent questionnaire. (2) The workshop provided the most frequent locale for meeting the friends gifted children made. One might say, therefore, that while the workshop, the room in which gifted children work together, helped to develop and reinforce friendships in-and-out-of-school, the regular class, which provides a place where gifted and typical children mingle and which is the really unique contribution of the Colfax Plan, did not actually produce relationships significant enough to be classified as friendships. This again calls attention to the fallacy of believing that "because we group children together we have trained them

to accept each other for what they are." Perhaps, if studies similar in methodology were done in complete segregation and in complete integration programs, a firm basis would be provided for general conclusions concerning the best provision for gifted children in our schools.

In an attempt to find the correlation between anxiety, intelligence, and achievement, McCandless and Castenada (1956) found that anxiety scales might be valuable in predicting achievement. If this correlation was consistent, the feeling of the students, especially the dull ones, might play an important part in the final results over a longer period of time not only in social adjustment but also in achievement.

Rudd (1958) attempted to measure attitudes, attainment, behavior, and personalities of the groups as influenced by ability grouping. The group included 180 pupils at the fifth grade level. He concluded:

> 1. There was very little difference in the results of the ability tests given.
> 2. The attitude toward school and school functions was relatively the same in both groups.
> 3. In ability-grouped classrooms there was less social contribution to lessons, more aggressive behavior, and less attention to work.
> 4. Teachers estimates of personalities showed no significant difference between groups. The pupils self estimates revealed an extensive but probably temporary deterioration in personality following regrouping.

Martin (1942) cited opinions of different individuals on the ability-grouping subject. These opinions either condemn the whole thesis of ability grouping or suggest extreme caution in adopting such a program.

Conant's study (1960) leads him to recommend a type of ability grouping. In his progress report on the junior high school, he says:

> In the fully departmentalized eighth grade, there should be ability grouping, preferably subject by subject in the areas of English, social studies, mathematics, and science. In this grouping there should be very few in both the top and bottom. Reading level tests as a major criteria for grouping may be preferable to I. Q. tests—they seem more relevant and are easier for the general public to understand.

In another quote he recommends:

> Interestingly, though grouping is a controversial subject, I have found considerably less objection to it in grades 7, 8, and 9 than I found three years ago in the senior high school. Many educators feel that by the time of the seventh grade the spread of pupil achievement has become so great that only an unusually competent teacher can provide suitable instruction for a cross-section of the grade. Complete homogeniety can never be attained, but a necessity is seen to reduce the range of individual differences in a given class if suitable instruction is to take place.
>
> I personally recommend three groups in academic courses with the bulk of the pupils in a particular grade in a large middle group. Preferably, the grouping should be accomplished subject by subject, except, of course, in those subjects combined in block-time classes. I have been especially impressed with the emphasis educators place on reading ability as one of the major criteria for grouping. Perhaps my principal argument for grouping in academic courses rests on the fact that in every school there is a certain fraction of pupils who read well below their grade level. These pupils need special books and teachers. To my mind, to mix in an English class boys and girls reading three years below grade level with those reading three years above grade level is to do everyone concerned an injustice. Of course, any grouping arrangement assumes differentiated materials and teaching methods.

Grouping and Classroom Range

Another problem involved in ability grouping is the contention that it is impossible to have a strictly homogeneous group. If we were to group students into three reading groups it would be necessary to regroup them for arithmetic. It is also contended that even though we started out with a homogeneous group, the differences in rate of maturity and other factors would soon produce a spread as great as the initial one.

From the files of the Department of Ungraded Classes in New York City, McElwee (1933) drew the records of 2,225 children. Their mental ages on the Binet scale ranged from six years to eight years and eleven months. Their reading and arithmetic scores were compared. The arithmetic achievement of 50 per cent of the entire group exceeded their reading achievement from two to six times. Fifty per cent of a homogeneously graded group based on reading achievement would be so heterogeneous in terms of arithmetic achievement that to suppose it was a homogeneous group would present serious difficulties.

Cook (1958), in his analysis of this problem, quotes Hull's study (1927). They feel that:

> Variability in the typical individual is 80 per cent as great as individual variability in his age group. Trait differences are normally distributed. Some individuals are twice as variable as others, and there is no relationship between general level of ability and of the amount of trait variability. Under favorable circumstances, that is, when pupils are grouped in x, y, and z fashion on the basis of an achievement test battery, which is heavily weighted in

> favor of reading and arithmetic scores, we may expect a reduction of about 20 per cent in reading and arithmetic variability . . . The extreme x and z groups will overlap approximately 80 per cent. Instead of a range of eight years in reading ability at the sixth grade level, the teacher has, after grouping, a range of six and four-tenths years. In other subjects such as art, music, handwriting, and spelling, the reduction of range approaches zero . . .

If Hull's findings are valid, then grouping for ability in the elementary school would present some difficult problems. It might be handled much easier in the secondary schools, however.

A study of grouping practices at a junior high school in Los Angeles by Ramey (1956) substantiates a trend just alluded to. These students were grouped on the bases of expectancy and reading ability. To these two criteria was added the recommendations of teachers and counselors.

It was found that there is no truly homogeneous group. Even though the class had been divided into four groups of 30 to 35 students each, they had to have groups within the group.

Within the range of test scores, there was almost complete overlapping of all groups except for the lowest.

From a report in the National Elementary Principal written by Cook (1958) the following statement is quoted:

> When a random group of six-year-olds enters the first grade, two per cent of them will be below the average four-year-olds in general mental development, and two per cent will be above the average eight-year-old. Disregarding the extreme two per cent at either end, there is a four-year range in general intelligence. By the time this

group has reached the age of twelve (seventh-grade level), the range will have increased to almost eight years. As long as all the children of all the people remain in school, the range continues to increase. When the educational achievement of a typical sixth-grade class is measured, we find a range of approximately eight years in reading comprehension, vocabulary, arithmetic reasoning, arithmetic computation, mechanics of English composition, and other forms of achievement. In almost any sixth-grade class there is a pupil with first—or second—grade reading ability, and another with eleventh—or twelfth-grade reading ability. In any grade above the primary level there is the complete range of elementary school achievement.

At the high school and college levels, Learned and Wood have given us an answer. When the General Culture Battery, consisting of achievement tests in general science, foreign literature, fine arts and social studies, was administered to high school and college seniors in Pennsylvania, it was found that the upper 10 per cent of high school seniors were above the college senior median and could have been given B. A. degrees without lowering the intellectual standards of such degrees. It was also found that the lower ten per cent of the college seniors were below the high school senior median.

Wrightstone (1957) concludes that:

Studies reveal that, in general, variability in achievement in grades that have three ability groups in each is about 83 per cent as great as in normally organized groups. In grades having two ability groups each, the variability in achievement, as measured by standard tests, is about 93 per cent as great as in normally organized groups. This difference offers only slight assistance to the teacher in reducing the range of individual differences in his classroom. For a grade organized on three ability levels, the reduction in range is about 15 to 17 per cent; for a grade with two ability groups, the reduction in range is 7 to 10 per cent.

Because there are wide differences even in a so-called ability group class and because it is difficult to avoid labeling classes as bright, average, or slow, homogeneous grouping has been less widely used in recent years than it was two decades ago. There have been

> developed both teaching methods and materials that permit more successful adaptations to a fairly wide range of ability within a class. Arguments have been advanced for and against heterogeneous grouping.

Certainly the literature does not possess any final answers for us. However, it suggests the complexity of the problem and issues even a greater challenge to explore the field for the answers that may be found.

CHAPTER III

METHOD OF PROCEDURE

Selection of Sample

This study involved fourth and sixth grade students in two adjoining school districts. The districts were classified district "A" (for the ability-grouped sample) and "R" (for the random-grouped sample) for the purposes of this study.

The experimental sample from district "A" consisted of 156 fourth grade boys, 132 fourth grade girls, 208 sixth grade boys, and 192 sixth grade girls.

These students had been placed in "developmental," "average," and "accelerated" classrooms within their grade level. Their placement was determined by their composite scores on the California Achievement Test Battery, Form W, with consideration given to teacher and principal evaluation.

The control sample from district "R" contained 173 fourth grade boys, 164 fourth grade girls, 261 sixth grade boys, and 222 sixth grade girls.

These students had been grouped at random on the basis of age and grade level.

All of the schools considered in the study were from comparable socio-economic areas. In other words, if a school in the experimental sample was from a low socio-economic area, a school in the control sample was chosen in the same area. If a school was chosen in a high socio-economic level area, one of the same type was chosen to control or match it.

Classrooms from which control subjects were chosen were selected on a random basis. The ability level for the children in district "R" was established on the basis of the Lorge-Thorndike Intelligence Test, Form A, the California Achievement Test Battery, Forms CC and AA, plus teacher recommendation.

Classification of sample

Early in the school year, students in both districts were given a pre-test using the California Achievement Test Battery as the measuring instrument. A team of people trained in educational testing administered the tests. A different form of the same test was given near the end of the school year.

Internal homogeneity scores based on the difference between the highest grade placement score earned by the pupil and the lowest grade placement score earned by the same pupil on the three different sections of the California Achievement Test (Battery 1) were calculated. Table 1 shows the distribution of internal homogeneity scores for fourth and sixth grade students in district "A" and district "R."

Table 1. Distributions of internal homogeneity scores, district "A" and district "R" fourth and sixth grade pupils, based upon California Achievement Test G-P scores on Battery 1

Internal homogeneity score (in months)	Frequencies					
	District "A"			District "R"		
	4th grade	6th grade	Total	4th grade	6th grade	Total
0		1	1		3	3
1	1	11	12	6	6	12
2	3	13	16	8	21	29
3	9	13	22	12	18	30
4	8	12	20	15	26	41
5	10	28	38	29	27	56
6	16	30	46	21	35	56
7	17	26	43	32	30	62
8	25	29	54	23	39	62
9	16	28	44	29	48	77
10	21	34	55	28	28	56
11	16	23	39	31	37	68
12	16	26	42	24	25	49
13	18	31	49	23	27	50
14	16	12	28	22	32	54
15	15	19	34	15	22	37
16	17	11	28	8	21	29
17	16	14	30	16	25	41
18	12	8	20	8	11	19
19	8	8	16	9	8	17
20	6	4	10	3	4	7
21	5	2	7	3	7	10

Table 1. Continued

Internal homogeneity score (in months)	Frequencies					
	District "A"			District "R"		
	4th grade	6th grade	Total	4th grade	6th grade	Total
22	4	1	5	2	11	13
23		3	3		5	5
24	1	3	4	1	3	4
25	3	4	7		3	3
26	1	2	3		3	3
27	1	1	2	1		1
28					1	1
29	1	2	3		1	1
30	1		1		1	1
31	2		2		1	1
32					1	1
33		1	1			
34					1	1
35						
36					1	1
37						
38					1	1
	N=285 M=12.05 SD= 5.65	N=410 M=10.01 SD= 5.75		N=379 M=9.85 SD=4.97	N=533 M=10.65 SD= 6.14	

After these distributions were made, mean and standard deviations for the two fourth grade samples and the two sixth grade samples were calculated. Cutoff points were then established so as to divide each grade level into three groups: (1) students of high heterogeneity, (2) regular students, and (3) students of high homogeneity.

The regular students were those that fell approximately + or - one standard deviation from the mean while those more homogeneous were placed in the homogeneous group and those more heterogeneous were placed in the heterogeneous group. These limits for the fourth grade were scores of zero to seven months for the homogeneous group, scores of eight to 15 months for the regular group, and scores of 16 to 31 months for the heterogeneous group. Cutoff points for the sixth grade were scores of zero to four months inclusive for the homogeneous group, five to 16 months inclusive for the regular group, and 17 to 38 months for the heterogeneous group. All distributions were skewed somewhat in the homogeneous direction. This was caused by a few heterogeneous cases being spread over a wide range. Table 2 shows the original tally in each district. Table 3 shows the final sampling used.

Statistical Procedure

After the above cutoff points were established, data were obtained from IBM so that a covariance analysis could be calculated comparing the achievement gains of homogeneous, regular, and heterogeneous pupils in the two districts and

Table 2. Original tally sheet[a]-total sample

Score variation	Number of students
District A	
1. .0-.4	85
2. .5-.8	193
3. .9-1.2	204
4. 1.3-1.6	137
5. 1.7-2.0	68
6. 2.1-2.4	20
7. 2.5-2.8	13
8. 2.9-Above	8
Total	728
District R	
1. .0-.4	112
2. .5-.8	248
3. .9-1.2	256
4. 1.3-1.6	179
5. 1.7-2.0	82
6. 2.1-2.4	29
7. 2.5-2.8	9
8. 2.9-Above	7
Total	922

[a]Each pupils three sub-scores on the California Achievement Test, on the 1st battery were compared. The difference in grade placement between the highest and lowest of these three scores was noted. This sheet gives a distribution of subjects in terms of the difference between highest and lowest grade placements scores on the three California Achievement Test sub-tests. A subject tallied under .0-.4 is internally homogeneous in terms of these scores.

Table 3. Final sampling

District	Homogeneous		Regular		Heterogeneous	
	Boys	Girls	Boys	Girls	Boys	Girls
			Fourth grade			
A	22	10	119	97	14	31
R	44	19	143	130	12	19
Total	66	29	262	227	26	50
			Sixth grade			
A	35	14	148	142	20	33
R	37	23	187	169	33	47
Total	72	37	335	311	53	80

at the two grade levels. It was hypothesized that heterogeneous pupils would achieve better under random-grouped system while homogeneous pupils would achieve better under the ability-grouped system. It was further hypothesized that no differences would occur between the two systems for regular students and that there would be no sex differences.

The three variables considered in this study were: (1) district, (2) level, and (3) sex. One variable was expressed while the other two were held constant. The covariance analysis was completed for each of the districts at each grade level. An interdistrict analysis was then completed at each level.

The results of these analyses were evaluated using the "T" Test and checked for significance both at the 5 percent level and at the 1 percent level. After the significance was determined, comparisons were made with the hypotheses and conclusions were drawn and summarized.

The results of this study offer possible answers for some of the many problems involved in the grouping process in education.

CHAPTER IV

RESULTS

Statistical Analysis

In order for the results of the analysis to be significant, it was found that a "T" Test of 2.57 was necessary to be significant at the 5 percent level and a "T" Test of 4.03 necessary at the 1 percent level.

This analysis will be concerned with the three variables: (1) sex, (2) type, and (3) district. Two of these variables have been held constant while the third one has been tested.

Sex Differences

To determine sex differences, comparisons were made as listed in Table 5. The hypotheses and the results of the analysis follow.

In the first hypothesis on sex, it was stated that boys are not significantly more homogeneous than girls. In the final tabulations the fourth grade sampling verified this with the total sample showing 7 percent boys and 5 percent girls. The sixth grade showed 7 percent boys and $4\frac{1}{2}$ percent girls.

Table 4. Mean score of internally homogeneous and heterogeneous pupils in district R and district A on California Achievement Test Batteries 1 and 2

District	Group	Battery 1 Mean	Battery 2 Mean
R	Homogeneous boys 4th	4.60	5.67
R	Regular boys 4th	4.43	5.38
R	Heterogeneous boys 4th	4.48	5.46
R	Homogeneous girls 4th	4.46	5.65
R	Regular girls 4th	4.68	5.81
R	Heterogeneous girls 4th	5.22	6.43
A	Homogeneous boys 4th	4.19	5.37
A	Regular boys 4th	4.12	5.29
A	Heterogeneous boys 4th	4.29	5.83
A	Homogeneous girls 4th	4.36	5.52
A	Regular girls 4th	4.64	6.11
A	Heterogeneous girls 4th	4.51	6.14
R	Homogeneous boys 6th	5.99	7.31
R	Regular boys 6th	6.02	7.27
R	Heterogeneous boys 6th	6.63	7.67
R	Homogeneous girls 6th	6.07	7.35
R	Regular girls 6th	6.21	7.50
R	Heterogeneous girls 6th	6.97	8.22
A	Homogeneous boys 6th	5.81	7.65
A	Regular boys 6th	5.69	7.28
A	Heterogeneous boys 6th	5.84	7.46
A	Homogeneous girls 6th	6.06	7.54
A	Regular girls 6th	6.05	7.75
A	Heterogeneous girls 6th	6.27	8.34
	Homogeneous mean	5.19	6.63
	Regular mean	5.23	6.55
	Heterogeneous mean	5.53	6.94

Table 5. District "A," fourth grade, sex and type differences, "T" test scores

	Boys			Girls		
	Homogeneous	Regular	Heterogeneous	Homogeneous	Regular	Heterogeneous
Boys homogeneous						
Boys regular	.07					
Boys heterogeneous	2.06	2.25				
Girls homogeneous	.24	.25	2.29			
Girls regular	1.64	1.78	.70	1.65		
Girls heterogeneous	2.72	2.92	.33	2.50	1.13	

Although both of these lean in favor of the boys, they do not reach the level of significance.

The second hypothesis on sex differences stated that boys at the various levels of homogeneity do not gain significantly more than do girls at the same level. The only significant difference on the fourth grade level was noted when regular boys and girls from district "R" were compared. This only approached significance at the 5 percent level.

In the sixth grade sampling, two comparisons approached significance at the 5 percent level. When district "A" sixth grade heterogeneous boys were compared with district "A" sixth grade heterogeneous girls and when district "R" sixth grade heterogeneous boys were compared with district "R" sixth grade heterogeneous girls, both difference approached significance at the 5 percent level.

It seems safe to conclude that a slight trend indicates that girls gain more than boys at the various levels, but nothing significant.

The next hypothesis deals with the number of homogeneous boys and girls at each level in each district. It is stated that there are not significantly more homogeneous boys in district "A" than in district "R."

This hypothesis was found to be supported at both the fourth and sixth grade levels in both districts. In the fourth grade sampling in district "A," 14 percent of the total sample was homogeneous boys. In district "R," 15

percent of the total sampling was homogeneous boys. This difference is far from significant.

The next hypothesis is the same as the preceding using the girls sample.

The fourth grade sampling here shows great differences in scores for fourth grade girls. District "A" had 12 percent of their girls in the homogeneous class while district "R" had only 7 percent. In the total range of scores, district "A" had only one half as many heterogeneous (11 percent) as did district "R" (22 percent). The sixth grade sample shows no significant difference with 8 percent for district "A" and 10 percent for district "R."

The conclusion must be made that there is a 14 percent difference in the number of homogeneous and heterogeneous girls in the fourth grade in the two districts, heterogeneous girls representing the greater number. As to what causes this, other than maturity patterns of girls, the answer is still open for speculation.

Type Differences

The purpose of this group of hypotheses was to see if a difference in achievement occurred among the three groups mentioned, viz., homogeneous, regular, heterogeneous.

It was hypothesized that students whose scores on the California Achievement Test Battery were homogeneous would achieve more than those that were regular or heterogeneous.

The results of the analysis showed no cases where homogeneous students proved superior to the other two types.

The reverse is indicated and will be discussed under the third hypothesis in this section.

The second hypothesis which suggested that regular cases would achieve at the slowest rate of the three types failed to show significance.

Hypothesis three revealed two areas of significance at both the fourth and sixth grade levels. In the fourth grade sample the comparisons between homogeneous boys from district "A" and heterogeneous boys from the same district approached but didn't reach significance.

Two of the comparisons of significance were between district "A" fourth grade boys who were regular and fourth grade boys from the same district who were heterogeneous. The other was between fourth grade girls who were homogeneous and those that were heterogeneous, both from district "A." Both showed significance at the 5 percent level favoring the heterogeneous student.

The other comparisons which revealed significance were from district "A" also. One was the comparison between sixth grade homogeneous girls and sixth grade heterogeneous girls. This reached significance at the 5 percent level.

Comparisons between district "A" sixth grade regular and heterogeneous girls from the same grade and district again showed significance at the 5 percent level in favor of the heterogeneous sample.

Although only four out of the 24 comparisons proved significant, it does show in each of the cases that were

significant the advantage in favor of the heterogeneous student. Heterogeneous girls achieved more than regular girls in the sixth grade and heterogeneous girls achieved more than homogeneous girls also in the sixth grade.

Heterogeneous fourth grade boys achieved more than regular boys and heterogeneous fourth grade girls achieved more than homogeneous girls.

Another fact must be noted and that is that in the area of "type" the only significant differences were found in district "A." This leads us to our next comparison and that is interdistrict.

District Differences

One of the chief hypothesis to be tested in this study deals with achievement under the ability-grouping and the random-grouping systems. To test all types in both districts, the first hypothesis was that students who were homogeneous in terms of their California Achievement Test Battery scores in the ability-grouped situation achieve significantly more than do homogeneous students in the random-grouped system. It was reasoned that the ability grouping would serve the student who was achieving near the same in all subjects more than those that were up in one subject and down in another.

This hypothesis did not hold up in the fourth grade sampling. At this grade level, no significant difference was found. This would indicate that the factors involved

in the two systems offer to a fourth grade student, regardless of the level of homogeneity, similar opportunity for achievement.

This pattern did not hold true in other interdistrict comparisons. In the differences between regular fourth grade students of the two districts, the significance was at the 1 percent level in favor of district "A," and between the heterogeneous fourth grade samples, it again showed a significance at the 1 percent level in favor of district "A."

The sixth grade sample was consistant in the same way. In comparisons between similar groups in each district, the differences favored district "A" in all cases significant at the 1 percent level.

Whatever caused the fourth grade sample to be insignificant doesn't carry over to the sixth grade. It is noted with interest that of the three sixth grade comparisons, the homogeneous sampling had the lowest score even though it was significant at the 1 percent level. This tends to throw disfavor on the hypothesis that this type of student would do better in district "A."

The second hypothesis held that students who are regular would be about the same in both districts. It was concluded in the hypothesis that neither system would offer superior opportunities to this classification of student. The results however, showed significance beyond the 5 percent level favoring the ability-grouped district "A" program at the

fourth grade level. There is no apparent reason why the ability system should offer more to this student, yet the results are definite. The only possible answer comes from the newness of the program and the interest generated therefrom.

A similar trend was found at the sixth grade level but even more pronounced. District "A" regular pupils achieved more evidenced by a significance at the 1 percent level. This would indicate that even in the areas where one would assume no differences in instructional benefit, that the ability-grouping program offered advantages to its students.

In the heterogeneous classification where the greatest signs of progress and achievement are in evidence, we find even more dramatic differences.

It was hypothesized that students whose California Achievement Test Battery scores were heterogeneous would do better in the random-grouped situation. The results show the greatest differences between the two districts with this type of student. The results were significant beyond the 1 percent level in both the fourth grade sampling (4.58) and in the sixth grade sampling (9.60) in favor of district "A." This would strongly suggest that whatever causes a student to be heterogeneous in his California Achievement Test Battery scores finds fertile ground in the ability-grouped classroom. Certainly the hypothesis didn't hold up as expected.

There are several things that could account for the results of this calculation. If heterogeneous students, because of their one or two high scores, were placed in the accelerated classes, then the stimulation that comes from being in the highest group might account for some of the difference. Another speculation is in the area of the make-up of the test itself. If students were high in language and reading, by teaching a few advanced concepts in arithmetic, the score on the test raises significantly. It would appear that intelligence might play a large part in the calculations of this study. Although intelligence wasn't controlled, achievement was and the two are generally found to correlate.

Whether any of these suppositions are correct remains to be seen as further calculations are made in the parent study.

CHAPTER V

SUMMARY, FINDINGS, AND CONCLUSIONS

Summary and Findings

The purpose of this study was to evaluate what happens to students whose California Achievement Test Battery scores were homogeneous, regular, or heterogeneous when they are placed in an ability-grouped classroom as compared to those placed in a random-grouped situation.

The California Achievement Test Battery was administered near the beginning and also near the end of the school year.

According to their scores on the initial test, the students were classified as homogeneous, regular, or heterogeneous. That is, those whose scores were similar on the three different sections of the test, viz., reading, language, and arithmetic, were considered homogeneous. Those whose scores were greatly differentiated were classified as heterogeneous. Those whose scores were neither homogeneous or heterogeneous were classified as regular. Table 4 lists mean on both test batteries for each classification of student.

Analysis of covariance was then completed to test the following hypotheses with the results as listed:

I. There are no sex differences. In the study as a whole, it was hypothesized that sex differences would not be significant.

A. The study revealed that girls test scores on the California Achievement Test Battery were more homogeneous than were the boys scores in the same test battery, but the number was not great enough to be of consequence.

B. A slight trend was again noted favoring the girls when achievement scores were considered at each level, but nothing reached significance.

C. It was found that at the fourth grade level, district "A" had a larger percentage of girls in the homogeneous group (12 percent) than did district "R" (7 percent). It was also noted that district "R" had twice as many heterogeneous girls (22 percent) as did district "A" (11 percent). The other types showed no differences of note.

II. It was hypothesized that there would not be any differences between the various types of students in achievement gains made.

This hypothesis did not hold true. The heterogeneous groups showed more gains than did the others (see Tables 5 to 10). Heterogeneous girls achieved more than regular or homogeneous girls in the sixth grade significant at the 1 percent level. Heterogeneous boys achieved more than regular boys in the fourth grade significant at the 5 percent level and heterogeneous girls made more gains than homogeneous girls. This difference only approached the 5

Table 6. District "A," sixth grade, sex and type differences, "T" test scores

	Boys			Girls		
	Homogeneous	Regular	Heterogeneous	Homogeneous	Regular	Heterogeneous
Boys homogeneous						
Boys regular	1.67					
Boys heterogeneous	1.28	.21				
Girls homogeneous	.15	1.39	1.00			
Girls regular	1.91	.33	.00	1.22		
Girls heterogeneous	1.06	.25	2.17	3.05	3.25	

Table 7. District "R," fourth grade, sex and type differences, "T" test scores

	Boys			Girls		
	Homogeneous	Regular	Heterogeneous	Homogeneous	Regular	Heterogeneous
Boys homogeneous						
Boys regular	1.10					
Boys heterogeneous	.53	.21				
Girls homogeneous	1.08	2.40	1.40			
Girls regular	.60	2.12	1.00	.58		
Girls heterogeneous	.78	1.69	1.12	.14	.39	

Table 8. District "R," sixth grade, sex and type differences, "T" test scores

	Boys			Girls		
	Homogeneous	Regular	Heterogeneous	Homogeneous	Regular	Heterogeneous
Boys homogeneous						
Boys regular	.60					
Boys heterogeneous	1.14	.91				
Girls homogeneous	.13	.14	2.00			
Girls regular	.10	.88	2.81	.21		
Girls heterogeneous	1.90	1.40	2.08	.87	1.00	

Table 9. Interdistrict "T" test scores, sixth grade

	District "R" homogeneous	District "R" regular	District "R" heterogeneous
District "A" homogeneous	4.90		
District "A" regular	6.66	7.33	
District "A" heterogeneous	6.20	8.50	9.60

Table 10. Interdistrict "T" test scores, fourth grade

	District "R" homogeneous	District "R" regular	District "R" heterogeneous
District "A" homogeneous	1.13		
District "A" regular	2.75	3.63	
District "A" heterogeneous	5.00	5.70	4.58

percent level of significance. The only significant differences appeared in district "A" and they strengthened the findings just mentioned.

III. The third hypothesis pertained to differences between the two districts.

A. It was hypothesized that the homogeneous pupils would do better in the ability-grouped district. It was further hypothesized that heterogeneous pupils would do better in the random-grouped district and that regular students would do the same in both districts.

B. It was found that in all interdistrict comparisons except one, district "A" students achieved more than did students in district "R." Only in the homogeneous fourth grade sample did this trend break down. The results were as follows:

Fourth grade: Homogeneous students from district "A" as compared to homogeneous students from district "R" showed no significant difference. Regular students from district "A" as compared to regular students from district "R" showed significant differences at the 5 percent level. Heterogeneous students from district "A" as compared to heterogeneous students from district "R" showed significant differences at the 1 percent level.

Sixth grade: Homogeneous, regular, and heterogeneous students from district "A" as compared to homogeneous, regular, and heterogeneous students from district "R" all showed significant differences at the 1 percent level.

This countered the hypotheses that homogeneous students would be superior in district "A" and heterogeneous students would be superior in district "R."

Conclusions

Under systems comparable to the two educational programs considered in this study, one would be justified in expecting the following results:

1. Under the ability-grouped system, one could expect to find more homogeneous fourth grade girls. In the random-grouped district, one could expect to find more heterogeneous girls at the fourth grade level. The percent of difference was double (11 percent as compared to 22 percent).

Other than the differences mentioned above, all other sex differences are insignificant.

2. In the calculations pertaining to differences between types of students, heterogeneous classifications showed the only superior gains. It can be expected that heterogeneous students then make the best advancements in ability-grouped situations and make more progress than any other type.

3. When the two districts were compared, the students in the ability-grouped situation were found to be superior in their achievement gains. The greatest gains were made by the heterogeneous sixth graders, followed by regular sixth graders. Even the homogeneous sixth graders were significantly different to the 1 percent level.

The only place where this didn't hold was with homogeneous students in the fourth grade. No significant difference was found. No reason is evident for this.

4. In the total study, it is indicated that in the ability-grouped situation, on the basis of achievement, students do better. It doesn't appear, however, that it is differentially advantageous for students at various homogeneity levels.

LITERATURE CITED

Alberty, H. B., and O. G. Brim. 1936. The relation of the newer educational practices to grouping. The grouping of pupils. 35th Yearbook of the National Society for the Study of Education, Part I, p. 129.

Barthelmess, Harriett M., and Phillip A. Boyer. 1932. An evaluation of ability grouping. Journal of Educational Research 26:284-294.

Bateman, Grace, and A. F. Shultz. 1929. The mass instruction plan. Detroit Educational Bulletin 12, No. 8:18-19.

Billett, R. O. 1932. Provisions for individual differences, marking, and promotion. U. S. Office of Education Bulletin 17, Monograph 13.

Conant, James B. 1960. Education in the junior high school years. A Memorandum to American School Boards, Educational Testing Service, New Jersey.

Cook, W. W. 1958. Classroom methods, the gifted and retarded in historical perspective. Phi Delta Kappan 39:249-255.

Cornell, Ethel L. 1936. Effects of ability grouping determinable from published studies. The grouping of pupils. 35th Yearbook of the National Society for the Study of Education. Bloomington Illinois Public School Publishing Co., Part I, p. 290.

Courtis, S. A. 1923. Contributions of measurement. The Second Yearbook Bulletin of the Elementary School Principals, p. 161.

Hammond, Sarah Lou. 1959. Homogeneous grouping and educational results. Curriculum Letter. Department of School Services and Publications, Wesleyan University, Middletown, Conn.

Hartill, Rufus M. 1936. Homogeneous grouping. Bureau of Publications, Teachers College, Columbia University, New York City, New York.

Hull, Clark L. 1927. Variability in amount of different traits possessed by the individual. Journal of Educational Psychology 18:97-104.

Kefauver, G. N. 1929. Bases for ability grouping. Teachers College Record 31:99,113.

Keliher, Alice V. 1931. A critical study of homogeneous grouping. Contributions to Education No. 457, Teachers College, Columbia University, New York, p. 41.

Kepperling, Inez, Edith Shaw, and Grover Stout. 1929. The vertical grouping plan. Detroit Educational Bulletin 12, No. 7:4-5.

Kozal, L. J., and B. Z. Woods. 1958. What are ways of grouping students for effective instruction. National Association of Secondary School Principals Bulletin 42:16-20.

Kvaraceus, William C., and Marion E. Wiles. 1938. An experiment in grouping for effective learning. Elementary School Journal 39:264-268.

Luchins, A. S., and E. H. Luchins. 1948. Children's Attitudes toward homogeneous groupings. Pedagog Sem. 72:3-9.

Mann, H. 1957. How real are friendships of gifted and typical children in a program of partial segregation. Exceptional Child 23:199-201.

Martin, E. D. 1942. Ten years of ability grouping. Texas Outlook 26:28.

McCandless, Boyd R., and Alfred Castenada. 1956. Anxiety in children, school achievement and intelligence. Child Development 27:379-382.

McElwee, Edna W. 1933. Homogeneous grouping of retarded pupils. Elementary School Journal 34:118-121.

McGaughy, J. R. 1930. Homogeneous grouping of pupils. Childhood Education 6:291-296.

Miller, W. S., and A. J. Otto. 1930. Analysis of experimental studies in homogeneous grouping. Journal of Educational Research 21:95-102.

Nash, E. A. 1942. Project in ability grouping. Business Educational World 23:186-188.

Opinion poll. 1955. Age vs. ability grouping. Nation's Schools 56:6.

Ramey, Arthur G. 1956. A new look at ability grouping in the junior high school. California Journal of Secondary Education 31:289-291.

Rankin, Paul T. 1931. Pupil classification and grouping. Review of Educational Research 1:200-230, 238-244.

Rankin, Paul T., C. T. Anderson, and W. G. Bergman. 1931. Ability grouping in the Detroit individualization experiment. 35th Yearbook of the National Society for the Study of Education, Part I, p. 277.

Riley, F. C. 1956. Grouping gives each child a chance. Nation's Schools 58:51-55.

Roberts, I. L. 1947. Homogeneous grouping: Experiment. Chicago School Journal 29:30-32.

Rudd, W. G. A. 1958. The psychological effects of streaming by attainment. British Journal of Educational Psychology 28:47-60.

Theisen, W. W. 1922. The relative progress of VII-B groups sectioned on the basis of ability. Journal of Educational Research 5:295-305.

Turney, A. H. 1931. The status of ability grouping. Educational Administration and Supervision 17:21-42, 110-127.

Vreeland, W. 1930. Detroit's experiment on individualization. School and Society 32:398-402.

Wallin, J. E. W. 1928. Sectioning according to ability in 1941 and 1942. School and Society 56:526-529.

Washburn, Carleton W. 1924. The attainments of gifted children. 23rd Yearbook of the National Society for the Study of Education 23:247-261.

Wrightstone, J. Wayne. 1957. Class organization for instruction. Department of Classroom Teachers, National Education Association.

Wyndham, Harold S. 1934. Ability grouping. Educational Research Series, No. 31, Melbourne: Melbourne University, p. 156.