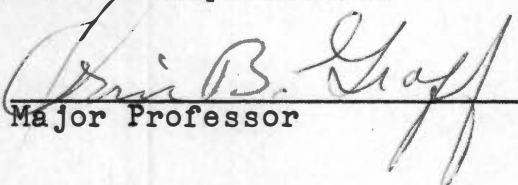


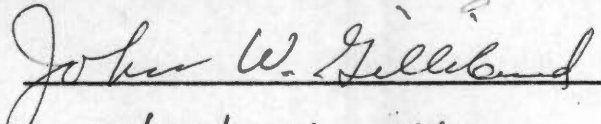
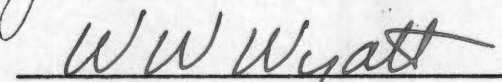
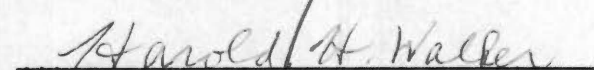
July 25, 1958

To the Graduate Council:

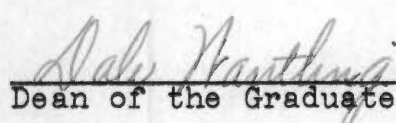
I am submitting herewith a thesis written by Helen Stratton Powers entitled "Changes in the Characteristics of Master's Students in Educational Administration." I recommend that it be accepted in partial fulfillment of the requirements for the degree of Doctor of Education, with a major in Educational Administration and Supervision.


Major Professor

We have read this thesis and
recommend its acceptance:

Accepted for the Council:


Dean of the Graduate School

CHANGES IN THE CHARACTERISTICS OF MASTER'S STUDENTS
IN EDUCATIONAL ADMINISTRATION

A THESIS

Submitted to
The Graduate Council
of
The University of Tennessee
in
Partial Fulfillment of the Requirements
for the degree of
Doctor of Education

by
Helen Stratton Powers
August 1958

ACKNOWLEDGEMENT

33

A list of those who helped make this study possible would begin with the writer's beginning. The wit and philosophy of her father, Walter Stratton, and the integrity and optimism of her mother, Ethel Mason Stratton, have been a constant source of support to the writer. Her late sister, Clara Belle Stratton, was always dependable for providing encouragement and sparkling humor.

During the months of graduate study, association with other students has stimulated and helped shape ideas." Former professors of the University of Tennessee who were especially helpful were Dr. Miles E. Cary, Dr. Robert S. Fleming, and Dr. Ralph B. Kimbrough. Each committee member was a good listener when the writer became besieged with problems which harass students. Committee members were Dr. Orin B. Graff, chairman, Dr. W. W. Wyatt, Dr. Harold H. Walker, Dr. Galen N. Drewry, and Dr. John W. Gilliland. Dr. E. Ohmer Milton, Psychology Department, the University of Tennessee, contributed suggestions as the writer developed the study.

The manuscript was improved by criticisms by Anne Hancock Farris and William R. McNabb. Mrs. Harold Wimberly's skill and patience took care of the preparation of the copy for reproduction purposes. William D. Powers, the writer's husband, helped in at least one hundred different ways. His cooperation and assistance made the study possible at this time. The writer thanks all the people, named and unnamed, who helped make this study a reality.

411373

TABLE OF CONTENTS

CHAPTER	PAGE
I. THE PROBLEM	1
General introduction	1
Statement of the problem	5
Importance of the study	7
Definition of terms	9
Scope of the study	10
Methods of procedure and sources of data . .	11
Organization of the study	14
II. DESCRIPTION OF THE EXPERIMENTAL PROGRAM . . .	16
Purpose of the experiment	16
Composition of the 1957 summer group . . .	19
Activities that implemented the program . .	23
Immediate outcomes	31
Chapter summary	40
III. CHANGES IN THE PARTICIPANTS AS A GROUP AND	
A COMPARISON WITH CHANGES REPORTED IN	
AN EARLIER EXPERIMENTAL GROUP	41
Group change by Rating Guide categories . .	43
Group change by Rating Guide sections . .	54
A comparison of change in this group with	
change in an earlier group as indicated by	
the Tennessee Rating Guide	59

CHAPTER

PAGE

III. (continued)	
Chapter summary	64
IV. DESCRIPTION OF THE GROUP AS INDICATED BY TESTS, AND DIFFERENCES BETWEEN INDIVIDUALS WHO CHANGED AND THOSE WHO DID NOT CHANGE . . .	65
Test results	66
Change as indicated by the Rating Guide . .	78
Correlation of change and test results . .	80
Important differences	85
Chapter summary	93
V. A STUDY OF THE 1956 PILOT GROUP TO NOTE EFFECT ONE YEAR LATER OF EXPERIMENTAL PROGRAM . .	95
Composition of the pilot study group . .	95
Individual ratings	96
Group ratings	98
Ratings by Rating Guide Sections	98
General evaluation	100
Chapter summary	102
VI. FINDINGS, CONCLUSIONS, AND RECOMMENDATIONS .	103
Summary of findings	104
Conclusions	109
Recommendations	112
BIBLIOGRAPHY	115
APPENDIX	118

LIST OF TABLES

TABLE	PAGE
I. Tabulation of Class Concerns as Indicated by Diaries	33
II. Summary of Class Evaluation of Diary Use . .	34
III. Pre-rating and Post-rating Averages of Individuals in the 1957 Experimental Group . .	45
IV. Pre-rating and Post-rating Averages for 1957 Group by Rating Guide Categories . . .	47
V. The Average Scores of the 1957 Experimental Group on the Various Sections of the Rating Guide as Indicated by Pre-rating and Post-rating	55
VI. Numbers and Percentages Showing Change by Sections of the Rating Guide as Indicated by Pre-ratings and Post-ratings of the 1957 Experimental Group	57
VII. Mean Scores of the 1956 and 1957 Groups for the Allport-Vernon Study of Values	68
VIII. Mean Scores of the 1956 and 1957 Groups for the Edwards Personal Preference Schedule	70
IX. Mean Scores of the 1956 and 1957 Groups for the Thurstone Temperament Schedule	73

TABLE

PAGE

X.	Upper Group Scores for Certain Tests	89
XI.	Lower Group Scores for Certain Tests	90
XII.	Ratings of Individuals in 1956 Group for the Three Years, 1956, 1957 and 1958	97
XIII.	1956 Group Average by Rating Guide Sections for 1956, 1957, and 1958	99

LIST OF FIGURES

FIGURE		PAGE
1.	Composition of the 1957 Summer Experimental Group, Showing Distribution of Twenty-One Students	21
2.	Rating Guide Averages for Individuals in the 1957 Experimental Group, Indicating Change Between 1957 Pre-rating and 1958 Post-rating	44
3.	Averages for 1957 Group by Rating Guide Category, Showing 1957 Pre-rating and 1958 Post-rating	46
4.	Group Change Between First and Second Year Ratings for Each Rating Guide Category for Both the 1956 and the 1957 Groups	63
5.	Rating Guide Averages of Individuals in Both the 1956 and the 1957 Groups Showing Change Between Pre-rating and Post-rating for Each Individual, Showing Upper and Lower Thirds of Group	79
6.	Graphs Suggesting Positive Correlation Between Great Amount of Change Between Pre-rating and Post-rating and High Scores on Certain Tests, a. <u>dominance</u> , Edwards, b. Minnesota Teacher Attitude, c. Watson-Glaser Critical Thinking	82

FIGURE

7. Graphs Suggesting Positive Correlation Between Great Amount of Change Between Pre-rating and Post-rating and High Scores on Certain Tests, a. active, Thurstone, b. dominant, Thurstone, c. stable, Thurstone 83
8. Graphs Suggesting Negative Correlation Between Great Amount of Change Between Pre-rating and Post-rating and High Scores on Certain Tests, a. succorance, Edwards, b. abasement, Edwards, c. P, PMS Scale, d. M, PMS Scale 84
9. Scale Showing Test Range of Combined 1956 and 1957 Groups for Ten Tests, dominance, Edwards; Minnesota Teacher Attitude; Watson-Glaser Critical Thinking; active, dominant, stable, Thurstone; succorance, abasement, Edwards; P, M, PMS Scale (Mean Score Profiles are Indicated for Upper and Lower Thirds) 86
10. Samples of Individual Profiles versus Upper Group Profile Based on Tests Shown in Figure 9 (Individuals A and C Represent Upper Third; Individuals B and D, Lower Third) 87

CHAPTER I

THE PROBLEM

General Introduction

In 1951 the Department of Educational Administration and Supervision of the University of Tennessee in cooperation with the Southern States Cooperative Program in Educational Administration (succeeded in 1956 by the Associated Programs in Educational Administration) initiated a research project designed to improve the preparation program of the Department and to provide a basis for improving the selection of students. The research staff at the University of Tennessee has largely centered its activities around an attempt to define the characteristics needed for effective educational administration.

Characteristics which would distinguish between effective and ineffective administrative performance were identified and stated in behavioral terms early in the project. The statements were incorporated in a single instrument which was termed "Tennessee Rating Guide."¹ The intended use for the Rating Guide was that it serve as a criterion in the selection of potential educational administrators.

¹Orin B. Graff and Ralph B. Kimbrough, "What We Have Learned About Selection," Phi Delta Kappan, 38:295, April 1956.

Several studies were undertaken to establish the validity of the Rating Guide. For example, behavioral characteristics of practicing administrators were observed, in separate studies, by Greever², Gentry³, and Schmitt.⁴ The original Rating Guide was revised from time to time as research indicated need for change. The January 1958 revised form is included in the Appendix of this study.

Other studies were made to determine correlations between psychological tests and the Rating Guide (in sections or as a whole). Such tests would be more economical and more convenient to employ than the Rating Guide. Moffett⁵, Luton⁶,

²Clarence E. Greever, "A Study of the Characteristics of Selected Effective Superintendencies in East Tennessee" (Unpublished Ed. D. thesis, The University of Tennessee, June 1956).

³Harold Wayne Gentry, "Patterns of Behavioral Characteristics Exhibited by School Administrators" (Unpublished Ed. D. thesis, The University of Tennessee, March 1957).

⁴Leonard Schmitt, "The Behavioral Characteristics of School Principals" (Unpublished Ed. D. thesis, The University of Tennessee, June 1958).

⁵Charles Moffett, "Operational Characteristics of Beginning Masters' Students in Educational Administration and Supervision" (Unpublished Ed. D. thesis, The University of Tennessee, June 1954).

⁶James N. Luton, "A Study of the Use of Certain Standardized Tests in the Selection of Potential Administrators" (Unpublished Ed. D. thesis, The University of Tennessee, March 1955).

and Nunnery⁷, in separate studies, found correlations but did not find adequate significance to permit the substitution of psychological tests for the Rating Guide.

After evidence indicated that the Rating Guide would distinguish between effective and ineffective administrators (in terms of the theory of administration subscribed to in the Rating Guide) it was feasible to attempt to determine how to change persons from ineffective to effective prospective or practicing administrators through a preparation program.

An experimental program was designed for the summer of 1956 at the University of Tennessee. The seventeen individuals who participated as subjects in the experimentation were beginning Master's students. The program was based on two assumptions: (1) as a prerequisite to effective behavior a person must understand his present characteristics in relation to the demands for effective job performance; and (2) an individual is more likely to change his behavioral characteristics through the process of self-understanding if he is placed in a democratic, group problem-solving climate for learning where the solutions of problems demand understandings by the student of himself, of his roles and his relationships, as

⁷Michael Yates Nunnery, "A Study in the Use of Psychological Tests in Determining Effectiveness and Ineffectiveness in School Administrators" (Unpublished Ed. D. thesis, The University of Tennessee, June 1958).

well as of the external problem situation of which he is a part.⁸

The program was designed within a framework of "blocks of time" of at least three hours. The time spent thus in a group situation contributed to a concentration of effort. Experiences succeeding experiences without the intervention of unrelated demands provided the opportunity for a mental and psychological climactic effect, producing the opportunity for a high degree of reflection, insight or expression. Part of any class period is generally spent in the recapitulation of experiences or plans of the previous period. Short, intermittent, fifty-minute periods which typify class scheduling would have required that group achievement be attained quickly, or be postponed or lost, during the portion of the period that would remain.

An important implementation to the program was the fact that the teaching staff was dual in nature; that is, it included a professor of psychology (who was also a clinical psychologist) and a professor of educational administration and supervision. Two distinct contributions were made by the dual-natured staff: (1) knowledge in the areas of psychology and educational administration was integrated for the students, and

⁸Gem Kate Taylor, "The Development of Effective Characteristics in Students of School Administration" (Unpublished Ed. D. thesis, The University of Tennessee, June 1957), p. 6.

(2) the students were able to witness an exchange of attacks and rebuttals related to educational or psychological issues between the professors. The exchange served to illustrate that ego-involvement is a barrier that needs to be eliminated prior to growth in the area of problem-solving.

The outcome of the program was that the experiment was considered successful as a pilot study; as a group the individuals showed certain desirable changes in behavioral characteristics. It was recommended that another such program be undertaken to help eliminate the possibility that the first group was non-representative of the general population of prospective administrators in education. Also, because of the wide variations among individuals a larger number of persons involved in such a program would furnish a better sample size for evaluative purposes.

Statement of the Problem

The problem in this study was to describe and report the effect of a second experimental program which was conducted in the summer of 1957 by the Department of Educational Administration and Supervision. As indicated above, the 1957 program was undertaken to observe the extent to which the group of students involved in this program behaved like the group involved in the pilot study one year earlier.⁹

⁹Ibid.

In the 1956 study an analysis had been made to determine whether certain behavioral characteristics could be developed in prospective educational administrators through increased understanding of one's personal characteristics in relation to his job. The 1957 summer experiment was the same as the 1956 summer experiment with certain modifications. A description of the modified program was included in the present investigation. Some evaluative techniques used in the 1956 study were used in the present study; some were not used; and others were added. The present study, in addition to noting similarities and differences between the group participating in the previous study and the group in the present experimental program, proposed a list of identifiable personality traits that affect behavioral change, and suggested further steps in experimentation.

The hypothesis of the study was that methods could be employed for discerning and measuring changes in certain characteristics of beginning Master's students, and that distinguishing characteristics between those who change and those who do not change could be, at least grossly, determined.

Assumptions of the study were: (1) characteristics essential for effective administration had been stated in the Tennessee Rating Guide and were understood by the professors participating in the program; (2) instruments were available

that would assess certain skills and traits of individuals tested; also, qualified people for administering, scoring and interpreting the instruments were available; (3) the preparation program for graduate students in educational administration and supervision could be improved to more nearly meet the needs of the students; (4) criteria for determining growth, at least roughly quantitatively, could be established; and (5) students enrolled in the graduate program in educational administration and supervision would desire to develop effective characteristics and would attempt to grow in that direction.

Importance of the Study

As the conscious need for qualified educational administrators has increased, the departments of colleges and universities responsible for preparing the administrators have examined their program offerings. A changing society has created changing requirements for leadership. A person who accepts responsibility for educational administration must embrace certain fundamental concepts about individuals and society in order to be effective. Some of the values which affect the system of education in a democratic society may be stated as follows:

A foundational concept of the democratic faith is belief in the worth and dignity of the individual. Each individual is unique and is believed to

be able to make some contribution to society, regardless of his present condition. Therefore, each individual should be developed to the maximum of his capacities.

The essential equality of all individuals is also assumed in a democratic society. This belief certainly does not mean that people are equal in ability or that they should have equal rewards or possessions. However, it does mean equality before the law, equality in dignity as a human personality, equality or opportunity for education suited to the individuals' needs and abilities. It is assumed that with the recognition of such equality any person has opportunity to achieve whatever his ability merits.

. . . Perhaps one of the most fundamental characteristics of democracy is the manner in which it advocates broad participation in decision-making, in particular insuring participation of an individual in those decisions which directly affect him. This, fundamentally, is a belief that men have a right to govern themselves, that the state exists for the welfare of the individual and that government derives its power from the consent of the governed. It means establishing social goals and direction through consent and not through violence or imposition.¹⁰

Two questions have then become: What are the requirements for an effective educational administrator? Can educational leaders who may have been accustomed to a non-democratic educational society be changed?

At least a partial answer to the first question has been evolved and stated in the Tennessee Rating Guide. The Guide has been submitted to critical examination and has been modified within its framework of democratic concepts. The second question was given consideration in the summer of 1956 at the University of Tennessee in a pilot study. The study

¹⁰Southern States Cooperative Program in School Administration, Better Teaching in School Administration (Nashville: SSCPEA, 1955), p. 118.

indicated that individuals can change. An underlying thesis was that change within an individual must begin with an understanding of himself and with his subsequent desire to change rather than for change to be imposed by an outsider.

The individuals, as a group, in the pilot study did manifest observable behavioral changes, as indicated by Rating Guide measurements, within the period of the year in which they participated as experimental subjects. Since some individuals changed and others did not change other questions arose: Are there any identifiable common personality factors that affect change? Do some people change more slowly than others? Are apparent changes actually transient in nature? The present investigation was undertaken to test findings of the 1956 pilot study and to seek answers to the above questions.

Definition of Terms

A behavioral characteristic. A characteristic identified through observation of behavior. The items included in the Rating Guide (see Appendix) describe behavioral characteristics.

A personality trait. A relatively permanent characteristic that can be identified by response to psychological tests.

A rating. The location on the Rating Guide of an

individual after that person has been observed in the performance of his stated job.

A pre-rating. A rating made of an individual prior to his participation in an experimental educational program in the Department of Educational Administration and Supervision.

A post-rating. A rating made of an individual at some time following his participation as a student in an experimental educational program in the Department of Educational Administration and Supervision.

A rating team. Usually two persons, either graduate students and/or professors, who made field visitations for the purpose of rating the students who were involved as subjects in the summer experimental programs.

Scope of the Study

In an effort to establish some critical factors pertaining to change in individuals this study dealt with the problem in four ways. One was a description of the experimental program of the summer of 1957. Included was an account of the composition of the participating group and of the design and implementation of the program itself. Another was a simple comparison of outcome between the 1957 and the 1956 experiments. The comparison was based on pre-ratings and post-ratings of members in both experimental groups.

A third portion of the study was an examination of data about the individuals in both groups combined. These data were constituted largely of psychological test scores compared with change indicated by the Rating Guide. The fourth part was a report of ratings of the 1956 subjects one year following the completion of the pilot study in which they had been involved.

Other studies in progress will describe a control group that was set up in the summer of 1958 for the purpose of further validating findings of the summer experiments of 1956 and 1957. The effects of a 1958 summer program for post-Master's students designed in the same manner as the earlier studies will be examined.

Methods of Procedure and Sources of Data

In the early part of 1957 the Department of Educational Administration and Supervision of the University of Tennessee sent letters to principals in the radius from which graduate students generally come to the University. The letters had requested the cooperation of the principals in notifying potential candidates for the Master's degree in educational administration and supervision about the 1957 experimental program plans. Interested persons were invited to complete an application to the program and were requested to obtain approval of their administrative superior for a visit to them

in their schools by research representatives from the Department. Applications were received and visits were made. One weakness of the procedure was the expense risked in visits to people who were uncertain about their graduate school plans. Many persons whose plans were tentative failed to enroll in summer school.

Ratings were made following visitations to the applicants. Not enough persons had indicated an interest in the experimental program by the time of summer school registration to compose a reasonable-sized group. At registration, graduate school staff members counselled with beginning students to encourage those who planned to attend both summer terms to participate in the program. A total of twenty-one persons was enrolled in the program.

Those who had not been visited and rated were asked to furnish names of administrative superiors or of staff colleagues who were in close proximity to the University. One or more of these persons were interviewed in relation to each previously unrated student and were asked to complete a rating of the student he knew. The pre-ratings were to be used later for comparison with post-ratings.

In early class meetings psychological tests were administered to the group. The tests dealt with attitudes, interests, critical thinking ability, and beliefs. Intelligence tests with performance and verbal scores were given, also. The tests were given to benefit the students and for

research purposes. The scored tests were returned to each individual and the tests were explained by the psychology professor for the purpose of increasing self-understanding in the student; also, the scores were coded and filed by a member of the research staff for later analysis.

In addition to tests, or inventories as many were properly termed, the meetings of the group were developed as a combination of a psychology course and an educational administration course. The psychology portion of the curriculum emphasized the development of personality and interpersonal relations. The education emphasis was an introduction to theory and practice in educational administration. Near the end of the summer classes, some of the tests were re-administered to aid further student self-assessment and to add to research data.

In the winter and spring of 1958 research teams made visits for rating purposes. All participants who were willing to continue as subjects in the experimentation and who lived within a 150-mile radius of the University were visited and rated. Eighteen of the twenty-one were included. No subject was rated more than once by the same rating team.

Letters were written to the members of the 1956 study group who were still in the education profession and who were within a 150-mile radius of the University. They were asked to permit a 1958 visit by research members of the Department.

Fifteen of the original seventeen were subsequently visited and rated. Ratings for individuals were compared with earlier ratings for evidence of change.

The test data of both groups were examined for evidences of differences between individuals who changed and those who did not change between ratings. Taylor's raw data were contributed to this study for these analyses.¹¹ Methods and techniques for processing data pertaining to social relations and changes in group situations were suggested by Jahoda¹² and Taba.¹³

Organization of the Study

The problem of this study is given in Chapter I. General introduction, statement of the problem, importance and scope of the study, methods of procedure, sources of data and organization of the study are given.

Chapter II states the purpose, method, and immediate outcomes of the 1957 summer experimental program. The composition of the group is described.

¹¹Taylor, op. cit.

¹²Marie Jahoda, Morton Deutsch and Stuart W. Cook, Research Methods in Social Relations, Part One: Basic Processes (New York: The Dryden Press, 1951), pp. 95-100.

¹³Hilda Taba et al., Diagnosing Human Relations Needs (Washington: American Council on Education, 1955), p. 5.

In Chapter III, changes observed in the 1957 experimental group are reported and a comparison is made with changes that had been previously reported for the 1956 experimental group.

Characteristics, as indicated by tests, common to members of both groups who changed are listed in Chapter IV.

An analysis of the effect of the 1956 pilot program after a lapse of one year is made in Chapter V. Changes, as indicated by Rating Guide measurements, that had occurred in the year following the original study are noted.

A summary of findings, a list of conclusions and a list of recommendations compose Chapter VI.

CHAPTER II

DESCRIPTION OF THE EXPERIMENTAL PROGRAM

The basic components of the 1957 experimental program described in this chapter were purposes, methods, and immediate outcomes of the program, as well as the composition of the experimental group.

Purpose of the Experiment

One of the steps in the direction of the long-range goal of preparation program improvement was the 1957 summer experiment. Concern about the improvement of the preparation program in educational administration and supervision was discussed in the first chapter of this study. A pilot study had indicated the feasibility of the next step. The specific purpose of the 1957 experiment was two-fold: (1) to acquire additional evidence in relation to the belief that beginning Master's students can change behaviorally in a certain kind of educational setting (one that permits a non-censuring self-examination) and (2) to determine if a cluster of definable differences exists between those who change and those who do not change as a result of the educational experience.

Establishment of Certain Understandings

The staff members apprised the group of the purpose of the experiment, but did not tell them that they were being

subjected to ratings. The group was assured that the psychological test results would be kept confidential and would in no way affect their status as students. As a part of the pact between students and staff anonymity was guaranteed in any reports that would be made about the experiment. The students were urged to respond as accurately as possible in the testing; it was emphasized that the results were to be used in self-appraisal rather than to be used for placement or for staff evaluation.

In accord with educational theory which respects personality, the staff members could not and did not use coercion in an attempt to bring about change. The staff was conscious of the seeming dichotomy of the existence of significant individual differences and of the existence of much similarity of human needs. Staff members acted on the theory that

Purposeful motivated behavior is found as the fulfillment of personal needs is sought. The individual acts on the basis of his perception (of self and reality) and on the basis of subsequent satisfaction which he receives from his experiences.¹

Even with the intent of providing a situation which would permit the student to assess himself, the classes were not intended to be therapeutic or advisory in nature. No emphasis was placed on "group-ness" although group consciousness did gradually emerge. Alongside the emphasis on

¹Southern States Cooperative Program in Educational Administration, Better Teaching in School Administration (Nashville: SSCPEA, 1955), p. 117.

increased understanding of one's personality the nature and function of educational administration were presented.

The Rating Guide as a Measuring Device

The Rating Guide is an ordinal scale which describes the behavior of potential or practicing administrators who function in various capacities in the educational positions they hold. An ordinal scale defines relative positions of characteristics, behavioral characteristics in this case, with no implications as to the distances between positions. Jahoda describes an ordinal scale in the following graphic manner:

An ordinal scale is like an elastic yardstick which is being stretched unevenly; the scale positions as indicated by the numbers on the stick are in a clearly defined order, but the numbers do not provide any definite indication of the distance between any two points on the stick.²

Several persons interested in the Rating Guide have suggested that the Guide be developed into an interval scale. Such a scale would be convenient because of "the statistical refinement which (it would permit) in the analysis of data."³ Jahoda says this aspiration has been "rarely attained" in dealing with measurements of human attributes. Attempts have

²Marie Jahoda, Morton Deutsch and Stuart W. Cook, Research Methods in Social Relations, Part One: Basic Processes (New York: The Dryden Press, 1951), p. 121.

³Ibid., p. 125.

been and are being made to adjust the Rating Guide scale to such a degree that increased refinements of statistical treatment can be applied in its use. In the meantime the Guide has been subjected to validation and reliability studies. The most recent reliability study was completed by Evernden who conducted a rate-rerate experiment. Evernden states:

The reliability coefficients for the entire sampling of sixty principals rated by eight raters were significantly high. . . . It was concluded that the Tennessee Rating Guide for School Administrators had a high degree of statistical reliability when used by some raters, but that not all raters were capable of using it with the same degree of reliability.⁴

Composition of the 1957 Summer Group

Professional Experience of the Students

Members of the group who were experimental subjects showed a wide range in professional experience. Two in the group were inexperienced in the profession of education; two had had one year's experience. Years of experience ranged up to eighteen. Four members were elementary teachers, two were junior high school teachers, and nine were high school teachers. Four principals, two in schools of grades one through eight, one of grades one through twelve, and one of grades

⁴William Lyle Evernden, "The Reliability of the Tennessee Rating Guide for School Administrators" (Unpublished Master's thesis, The University of Tennessee, June 1958), p. 40.

nine through twelve, were in the group.

Although the group was composed principally of beginning Master's students, three of the four principals who were in the group had had some graduate study. (The ratings and test scores of those principals who had completed some graduate study scattered throughout the range of the ratings and scores of the entire group.)

Most of the individuals planned to continue in their positions of the preceding school year; two planned to change positions and one of the two who had had no experience planned to accept a high school teaching position. The other planned to continue graduate study.

Social Nature of the Group

Sixteen of the students were married. Eight of them were parents. Two members were a husband-and-wife combination. None had dependents in their homes other than children or a spouse. Six of the group were women. The group was bi-racial: two members were Negroes. Figure 1, page 21, illustrates the composition of the group.

The Staff

The summer school session was divided into two terms. The teaching staff was composed of a psychology professor and a professor of educational administration and supervision. Those who taught during the first term had conducted the

No experience	Teachers, Elementary	Teachers, High School (Junior and Senior)	Principals
Women		Men	
Single		Married	Parents
Negro	White		

Figure 1

Composition of the 1957 Summer Experimental Group,
Showing Distribution of Twenty-One Students

pilot study which had preceded the present experiment. At the end of the first term a prearranged change in the staff was made; a different professor of educational administration taught. This professor had visited the classes from time to time and had become acquainted with the group members and with the progress of the program. The three professors were married men, with children. They were regular staff members of the University of Tennessee. The professors of educational administration and supervision had had public school experience as teachers and as principals.

Two doctoral students served in the capacities of observer-assistant and observer-participant. The first student administered most of the psychological tests, made one lecture to the group, and procured films and other materials for instructional purposes. The student did not attend all classes but seemed to be accepted by the group as a staff member. The second student participated as much as possible as a member of the group, attending all class sessions. The student assisted in schedule-making and record-keeping, plotted test results, and tried to serve as a liaison person between professional staff and students whenever such a need occurred. This person, too, seemed to be accepted by the group. Jahoda suggests that people "get used to observers as long as the behavior of the observers convinces the group members

that they are no threat."⁵ The observer-assistant was a married man, with no children; the observer-participant was a married woman with no children. Both persons had had teaching and administrative experience.

Activities that Implemented the Program

The aims of the staff when setting up the program structure were similar to those of the pilot program:

1. To help the prospective administrator develop a keen sense of the importance of operational beliefs of school administration as a guide to his action.
2. To help each student develop a consistent point of view of school administration.
3. To develop understandings and beliefs about the origins of adult behavioral characteristics.
4. To develop the ability to understand one's motivations for action and how they affect his way of working with people.
5. To develop the ability to apply sound democratic problem-solving procedures to job performance.
6. To develop an operational understanding of content, knowledge, and skills in the critical task areas of school administration.
7. To develop an operational understanding of the role of the school administrator in the community.⁶

⁵Marie Jahoda, Morton Deutsch and Stuart W. Cook, Research Methods in Social Relations, Part Two: Selected Techniques (New York: The Dryden Press, 1951), p. 534.

⁶Gem Kate Taylor, "The Development of Effective Characteristics in Students of School Administration" (Unpublished Ed. D. thesis, The University of Tennessee, June 1957), pp. 17-18.

All activities were scheduled during a daily four-hour morning period except recommended reading, preparation of one formal speech per person, and one session per person for the individually administered intelligence test.

Structure and Content

During the first eight days emphasis was on testing and test results. Tests were spaced carefully to avoid fatigue; however, by the seventh day there were some expressions of distaste for the testing. During test days all other activities were arranged around the test schedule. Following this testing period a schedule was set up with the understanding that flexibility existed within the framework, subject to group agreement.

The first five to ten minutes of each session were assigned to personal diary writing. This was a voluntary activity and not all persons wished to write daily.

The second activity was a series of lectures by the psychology professor on the subject of personality development. Emphasis was placed on the long-range effects of extremely early childhood influences. Lectures were replaced intermittently by films, book reviews, and group discussions.

After a short "break," emphasis was directed toward educational administration. The content centered around the critical tasks of educational administration, existing theories in education, problem-solving as a functional

philosophy, the relationship of the school to other organizations in the community, and the Tennessee Rating Guide. Activities included lectures by the professors of educational administration and supervision or guest speakers, and problem-solving techniques.

A "coffee break" was an important part of the schedule. Most of the students liked to go to the near-by Student Center for light refreshment and conversation. This period provided opportunities for increased personal acquaintance with other students and the professors, allowed grievances to be aired informally and freely, and gave added time for discussing ideas that had been introduced in class.

From the end of the coffee break until the end of the class period activities were scheduled in which the students participated in various ways. During the first term most activities were of a total-group nature. The exceptions to this were role-playing and speech-making. There were discussions based on panel presentations, films, or subjects that were considered to be critical to the area of educational administration. During the second term many of the activities were of a small-group nature. Representatives of the small groups would report a summary of proceedings to the entire group.

Use of Diaries

Whether to use diaries and how to use them are questions that produce varying answers. In the pilot study,

diaries had been kept daily or weekly chiefly as a log of reactions or new insights. The diaries were collected and processed by a staff member at the end of the term. The students had been advised of the intended use of the diaries. Their use was considered successful for the intended purpose.

Recognizing that adults are sophisticated enough to camouflage their feelings for the benefit of other people, a decision was made to ask the members of the 1957 group to keep diaries that they themselves would be asked to process in a specified manner. Class time was scheduled at the beginning of each day for diary entries. The purpose of the diaries was to promote insight into one's personality; it was expected that people would make entries about the things which were creating pressures upon them.

Taba defends the use of diaries and other "unscored measuring devices" in the following manner:

They are not tests to be scored with keys; they are methods of stimulating and summarizing responses. . . .

Of course, open-ended techniques raise the question of objectivity. If there are no predetermined keys . . . what is there to prevent subjective interpretation or even outright reading of fixed personal prejudices into the data? The danger of subjective bias is, of course, real. But equally real are the dangers inherent in objective tests; these are the dangers of limited coverage and atomization of interrelated behaviors. For program-building it seemed worth while to sacrifice some objectivity for methods that preserve some

of the relationships between concepts, feelings, motivations, and behaviors.⁷

Use of Psychological Films

One of the techniques used by the psychology professor may be termed a "'focused' group interview." Following the viewing of a film, which was the experience upon which attention was "focused," the group was asked questions by the professor which helped bring out the affective and value-laden implications of the subject's responses and to determine the personal significance of attitudes. In a "group interview" situation individuals were free to speak or to refrain from speaking. Responses were spontaneous, specific, and self-revealing. Although the professor was a clinical psychologist and was conscious of the nature of a therapeutic situation, the class situations were not designed to be "psychotherapeutic." And although no pathological treatment was involved, the particular use of the films furnished unusual opportunities for the examination of attitudes and beliefs, group-wise and individually. "Inherently this type of interview is more flexible and of course requires more skill on the part of the interviewer than do the more structured types."⁸

⁷Hilda Taba et al., Diagnosing Human Relations Needs (Washington: American Council on Education, 1955), p. 5.

⁸Jahoda, Part One, op. cit., pp. 175-176.

Observer Report

As an index to participation in total group discussion an observer had been instructed to record voluntary spoken participation by individual members. The record was kept during the entire time the group was in class for one day; the results were reported to the group at the end of the period. The observer was a person who had been involved in various activities in relation to the group. The fact that the data were being recorded generally passed unnoticed by members of the group. The use of observers in small face-to-face groups is commented upon by Jahoda:

Although the use of systematic observation is still relatively new as a research method in the social sciences, there have been enough experiences with the tool, and enough common problems among persons who have tried the method to permit anticipation of certain decisions which will have to be made prior to any investigation using research observers An observer may be trained to concentrate on data in only one (kind of relationship); or he may be asked to watch a number of them. Exactly how the observer performs, will depend upon the theory and the points which are seen as providing optimal data for the problem being studied.⁹

Pictorial Projective Technique

The psychology professor asked the class members to participate in the development of a psychological projective technique. The professor was attempting to develop a simplified pictorial technique to be used later in clinical

⁹Jahoda, Part Two, op. cit., p. 520.

analysis. When the technique had been described and they were promised an individual analysis the students were eager to participate. The process was individually administered, requiring short oral responses to a series of pictures.

The fact that pictorial techniques "have long been popular" in the study of social attitudes is observed by Jahoda who gives the reasons for the popular usage as follows:

The Thematic Apperception Test (T.A.T.) has been the stimulus for several ventures. The T.A.T. consists of a series of ambiguous pictures about each of which the subject is asked to tell a story. . . . It is assumed that in describing the characters depicted, . . . the subject indirectly tells something about himself. Any person in the story with whose actions the subject concerns himself, with whom the subject may be conceived as identifying, represents a medium through which the subject expresses his own inner tendencies. The thoughts, the feelings, the attitudes, the inhibitions, etc., expressed by the characters with whom he identifies provide clues to his own tendencies.¹⁰

Speech-Critique Exercise

One of the skills required of an educational administrator is the ability to deliver an organized talk in a formal group situation. The class was divided into two groups (the division was expedient because of the numbers involved) for speech-making by each member. A list of criteria for effective delivery had been given to the group. The talks were tape-recorded. Chairmanship for conducting the critique

¹⁰Jahoda, Part One, op. cit., pp. 215-216.

sessions was rotated among the listeners; each speaker was subjected to a critique of his performance by the group. Each member had the opportunity to schedule a listening time for hearing his own tape after he had heard the class critique. The topics for the talks were chosen by the individuals who gave each talk. A second speech on the same subject was originally planned in order to give an opportunity for immediate gain from criticisms, but this activity was later eliminated from the schedule by group agreement.

Testing

Tests were administered during the first days of the program. Time was scheduled soon after each test was scored for discussing the tests. The scored tests were returned to the individuals; graphs had been prepared that showed by code the placement of each person on each test in relation to other members of the class. A class mean score was shown for each test. Students were permitted time to theorize or rationalize or to ask questions as each test was discussed. This sort of "open" treatment of test results seemed to be somewhat startling at first; some individuals were embarrassed by their scores even though the scores were privately known. But the group soon appeared to relax and to become able to accept results.

At the end of the term two tests, one dealing with critical thinking and the other with teacher attitudes, were

re-administered. They were given in time for the results to be returned to the class.

Other Activities

Other teaching devices used were role-playing, small and large group discussions, case studies, lectures, audio-visual aids, "resource" speakers, and panels.

Immediate Outcomes

Diaries

The use of the diaries proved to be of questionable value. As time passed fewer people wrote in the diaries. When the diaries were processed by the individual writers, the fact that the discontinuance of writing had occurred was discussed. Two of the reasons given were that the diaries were not kept in prominence as a matter of importance by the staff, and "When others stopped writing I did, too." Perhaps the fact that the frequency of writing diminished had a positive significance, however. It is possible that increased acquaintance with the program and with members in the group served to reduce tension and that pressures created by outside forces were reduced upon arrival in class. Some of the members admitted frankly that they had not quite believed the statement made by a staff member to the effect that "diaries would not be collected and would not be read by any

person other than the writer"; therefore, they had been sparing with entries.

The structure which was furnished for processing the diaries grouped the entries under two concerns: people and things. The members were asked to tabulate the number of times entries dealt with members of their family, members of the class, and others who might have merited their interest. The other tabulated list included such personal matters as home, transportation, school, position, and finances. Table I is a summary of the tabulations reported by the class. The purpose of the processing was to permit the individual another kind of glimpse at himself in order to see if there were some things about which he worried of which he may not have been previously aware.

When requested to evaluate the use of the diaries the majority felt they had not been particularly beneficial. Some appreciated the opportunity to use them as memorandum books; some felt that if the purpose of the diaries had been more fully explained their use might have been more helpful. Table II is a summarized list of the evaluation of the diary usage.

Psychological Films

The films and the ways they were used seemed to make quite an impact on the group. When people heard themselves stating prejudiced opinions aloud, for example, they began to

TABLE I

TABULATION OF CLASS CONCERNS AS INDICATED BY DIARIES

Area of concern	Frequency for item	Area of concern	Frequency for item
Family relations	25	Finance	15
Non-family relations	26	Housing	22
Job	28	Recreation	5
Summer School	68	Testing	11
Health	16	Personal	
Transportation	21	appearance	10

Note: Six other areas were listed by one person each.

Concern about myself	Frequency for item
What others think of me	27
Personal business	58
What I think about myself	25

Note: Two other concerns were listed by one person each.

Concern about others	Frequency for item
Family	
Spouse	44
Parent(s)	4
Older (or only) child	29
Younger child	7
Other family member	14
Summer school members	
Education professor, first term	22
Psychology professor	32
Education professor, second term	16
Student members	7

Note: Six other subjects were listed by one person each.

TABLE II

SUMMARY OF CLASS EVALUATION OF DIARY USE

Remarks about practical specific use	Number who gave the response
Primarily as a memorandum book	4
To combat tension or anxiety for groups or individuals	4
To be useful would have to consider motivation, stimulation, emphasis	3
Possible merit for homeroom or classroom situation	2
Mirror effect	2
No value	2
Helped to concentrate	1
Reasons suggested for discontinued use of diary	Number who gave the response
Lack of interest, no ends in view, no personal value	11
Time element	10
Influenced by other people	
Not told or reminded to write	4
Others did not regard it important	3
Childish	2
Recording of feelings either difficult or undesirable	2
Thought others might read it	1
Problems eased	1
Preferred to use time to talk	1

be able to evaluate the statements with a degree of objectivity. As individuals began to see how some of the fears and misunderstandings in their own lives are common to the lives of many other people they began to see themselves in a different perspective. Members became less defensive and more willing to laugh at themselves. This kind of growth was not total and did not proceed without frequent relapses. But, as a group, there was much improvement in self-insight as indicated by oral expressions.

Observer Report

The use of the observer-reporter (one of the doctoral students) to check group participation on one occasion was worthwhile for two reasons: (1) it demonstrated a device that could be adapted to group use in other situations, and (2) it pointed up by actual count that every member had volunteered at least one remark during the class period. One humorous outcome was that, unaware of the observer's role, two members had become involved in a fairly rapid exchange of words. As a result, one of them had spoken fifty times.

Pictorial Projective Technique

Interest was so high in learning about themselves that the members were eager to participate in a "trial run" of a projective technique being developed by the psychology professor. After making an analysis of the responses, however,

the professor reported that the project was a failure. Evidently the pictures had failed to be provocative enough to permit projective observations of a discriminating nature.

Speech-Critique

The assignment of speech-making seemed to be undertaken seriously. Criteria for effective speech-making, such as looking at the audience, avoiding the use of distracting personal gestures, and using illustrations, were discussed prior to the series of speeches. Topics were chosen on two bases: (1) they were to be subjects with which the speakers themselves were well acquainted, and (2) they were to pertain to the subject of education. Most of the students had already heard their voices on tape and all had had some experience in speaking to groups. But each person professed the experiencing of tension as he prepared to make his speech.

Most persons appeared to accept the criticisms of the group in the manner in which they were intended. Some off-the-cuff remarks suggested that a few people felt that some criticisms of themselves had been a little inappropriate.

The material chosen was interesting and informative, in most cases. The decision not to repeat the speeches was based on an estimate of the net gain from the expenditure of time involved. Most people appeared to be happy to be relieved of the responsibility of trying to improve the original speech; some expressed disappointment in having to give up the idea.

Testing

When the testing was begun there seemed to be apprehension in the group. Statements by the staff that the testing was "for the purpose of increasing self-understanding and not for grading" were received tentatively. When the "tests" proved to be "inventories" of attitudes, beliefs, and temperament rather than tests based on factual knowledge, fears seemed to subside. The final test in the initial series measured knowledge and skill in English mechanics and expression but, by that time, the test was approached with a minimum of dread, apparently. Test scores were recorded by code and were known to only one staff member (the doctoral student who was an observer-participant). The professors did not inquire about the scores of the individual students.

The results of each test were plotted on a graph that was posted on a bulletin board. The students seemed to look forward to locating themselves (in code) on the graphs. As one student observed, "One test will pick you up and the next one will let you down." Although there was occasionally some reaction of a negative nature, such feelings seemed to be short-lived. No demoralizing effect from the tests was in evidence as far as the total group was concerned.

Grades

The problem of grades for the course (actually the two terms included credit for four courses: two in psychology

and two in educational administration) was presented to the students as a "real" problem for their solution. During the first term solving their problem developed into a lengthy process. The solution was a check-list of items that indicated "readiness" for becoming an educational administrator. The check-list was scaled, with a wide range of possible scores. It was decided that every student in the group was to attempt to score every other member of the group. This rating or scoring took place near the end of the first term.

After the total scores were tallied, the results were placed on the chalkboard. Then, by group agreement, "natural" breaks in the scores were noted and grades were assigned on the basis of those divisions. At that point everybody knew his own grade because he knew his own total score.

Some tension and dissatisfaction had arisen in the group during the struggle to reach agreement about grading during the first term. The merit of such a struggle as a demonstration of "real" problem-solving was difficult to judge. The second term the students were the same (except for three who had left the group because of other commitments); the professor of educational administration was new but had attended several class meetings and knew the group. This professor immediately raised the question of grades for the second term. After a half dozen suggestions had been submitted to the class by its members, the suggestion that

each student grade himself at that time won approval. The reasoning was that the student had had time to comprehend the level and content of the course and had had time to judge the competence of the group as a whole. Also, he knew how much time and effort he expected to apply during the rest of the course.

Agenda

The professors presented several proposed activities to the group at the beginning of the first term. They explained which ones they considered essential to the course, and requested consideration of the others; they also solicited suggestions for additional appropriate activities. The kinds of activities which were ultimately rejected were those which involved visitation to observe such groups in action as a city council or a school board. The group felt that travel time would reduce the value of such undertakings.

Plans for the group were always made far enough in advance that there was a minimum of "lost motion" or uncertainty. There was no pretense or intention that this be an "unstructured" or "leaderless group" experiment.

Course Evaluation by Students

At the end of the second term the professor of educational administration and supervision asked the class members to submit unsigned written criticisms or suggestions

that might have improved the second term of the educational administration portion of the experimental program. Many students reviewed the program as a whole. The general design evidently had met the needs or had contributed to the consciousness of needs in a positive manner. One person summarized his comments in this way, "I don't think anything needs to be taken out; I would suggest only a few minor additions."

Other anonymous remarks which typify individual student reaction are quoted as follows:

I really think this particular course (small group work emphasizing problem-solving techniques) should be offered during the first term of the summer session. It would speed the forming of closer relationships and it would give more opportunity for democratic action.

I would suggest more work with problem solving-- as a whole, in small groups, or individually.

I can not find anything critically wrong with the program. In fact, I feel that the lectures, class discussions, and personal participation have been of extreme value to me. Also, it seems that one could receive a great deal from this program without being an administrator.

Chapter Summary

The experimental program in educational administration and supervision at the University of Tennessee in the summer of 1957 was described in this chapter. Included were a description of the composition of the group of students and staff, underlying purposes and theories for the experiment, the activities that made up the content of the program, and a report of the immediate outcomes of the activities.

CHAPTER III

CHANGES IN THE PARTICIPANTS AS A GROUP AND A COMPARISON WITH CHANGES REPORTED IN AN EARLIER EXPERIMENTAL GROUP

The Rating Guide scale ranges from one to five. Fractional ratings are not possible. Behavioral statements are made that range from one for the least desirable practice to five for the most desirable practice. The behavior that is the most nearly descriptive of the person under observation is indicated by checking the number (one, two, three, four, or five) that is beside it on the Rating Guide. (See Appendix of this study for a copy of the Rating Guide.) The order of the items in some categories of the Guide has been inverted to require alertness by the rater.

Rating Procedures

The students who were enrolled in the experimental program had all been rated by the first week of the 1957 summer school session. The rating is referred to in this study as the "pre-rating." A description of the method by which the students were pre-rated was given in the first chapter of this study. Twenty-one students were involved in the program. Distance or a change in interest caused three persons to be discontinued as subjects in the study. Change in those three individuals could not be measured; therefore,

all numbers in this chapter pertaining to the group are based on the remaining eighteen students.

The second rating, or post-rating, was made by persons who were not involved in the pre-rating of the individual. The post-ratings were made in the early part of 1958, at least eight months after the pre-ratings. Seventeen of the eighteen subjects were visited in their schools and were observed in the performance of their duties as teachers or principals. One person was a full-time graduate student. He was observed in classes and in other school activities.

The rating teams who made the post-ratings had not had access to the pre-ratings. The raters were graduate students and regular staff members of the Department of Educational Administration and Supervision; these persons were acquainted with the Rating Guide and were proficient in its use. The students who were visited in the field had been contacted by letter. Visits were arranged for times that were convenient to the visitants.

The students were gracious and courteous hosts. Rating team members were introduced to local students and staff members whenever feasible. The students seemed pleased to have an opportunity to discuss their local problems and plans with the visiting team; they also made inquiries about other members of the experimental program.

Differences in Pre-ratings and Post-ratings

Although each checked characteristic would be represented by a whole number, averages are represented by fractions. The average for the pre-rating of the group as a whole was 3.6. The post-rating average of the group as a whole was 3.9. Figure 2, page 44, shows the changes on a "total" Rating Guide basis for each individual. Pre-rating and post-rating averages are shown.

Table III is a tabular listing of the pre-rating and post-rating scores as indicated in Figure 2. Differences between the scores were obtained. An average change of 0.317 units and a variance of 0.1968 were obtained. These differences were analyzed by means of the t-test for paired differences and were found to be significant at the 0.05 level.

Group Change by Rating Guide Categories

Changes were described in terms of the Rating Guide. No attempt was made to justify changes. This portion of the chapter describes changes in the group as a whole. Paragraph headings are the titles of the various Rating Guide categories.

Figure 3, page 46, illustrates the 1957 and 1958 ratings of the 1957 group by category. Table IV is a tabular listing of the pre-rating and post-rating scores as indicated in Figure 3. Differences between the pre-rating and post-rating scores were obtained. These differences were analyzed

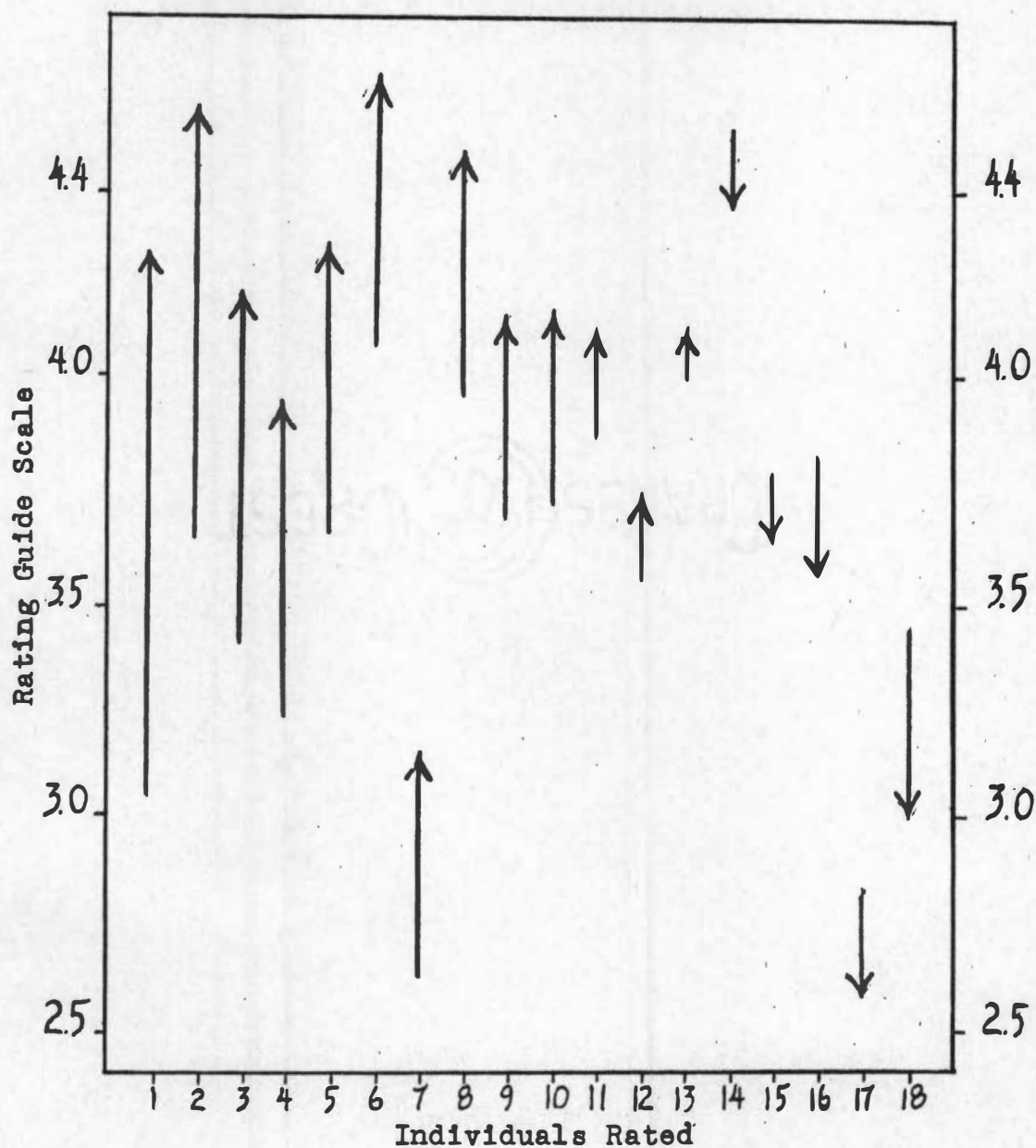


Figure 2

Rating Guide Averages for Individuals
in the 1957 Experimental Group, In-
dicating Change Between 1957 Pre-
rating and 1958 Post-rating

TABLE III

PRE-RATING AND POST-RATING AVERAGES FOR TOTAL RATING GUIDE OF
INDIVIDUALS IN THE 1957 EXPERIMENTAL GROUP

Individual	Pre- rating	Post- rating	Differ- ence
1	4.5	4.4	- 0.1
2	3.8	3.6	- 0.2
3	4.1	4.6	0.5
4	3.2	3.9	0.7
5	2.6	3.1	0.5
6	2.8	2.6	- 0.2
7	4.0	4.0	0.0
8	3.6	4.6	1.0
9	3.9	4.1	0.2
10	3.1	4.2	1.1
11	3.7	4.1	0.4
12	3.7	4.1	0.4
13	3.6	4.3	0.7
14	3.6	3.7	0.1
15	3.4	3.0	- 0.4
16	4.0	4.4	0.4
17	3.8	3.6	- 0.2
18	3.4	4.2	0.8
Total difference			5.7
Mean difference			0.317*

*Mean difference was significant at the 0.05 level.

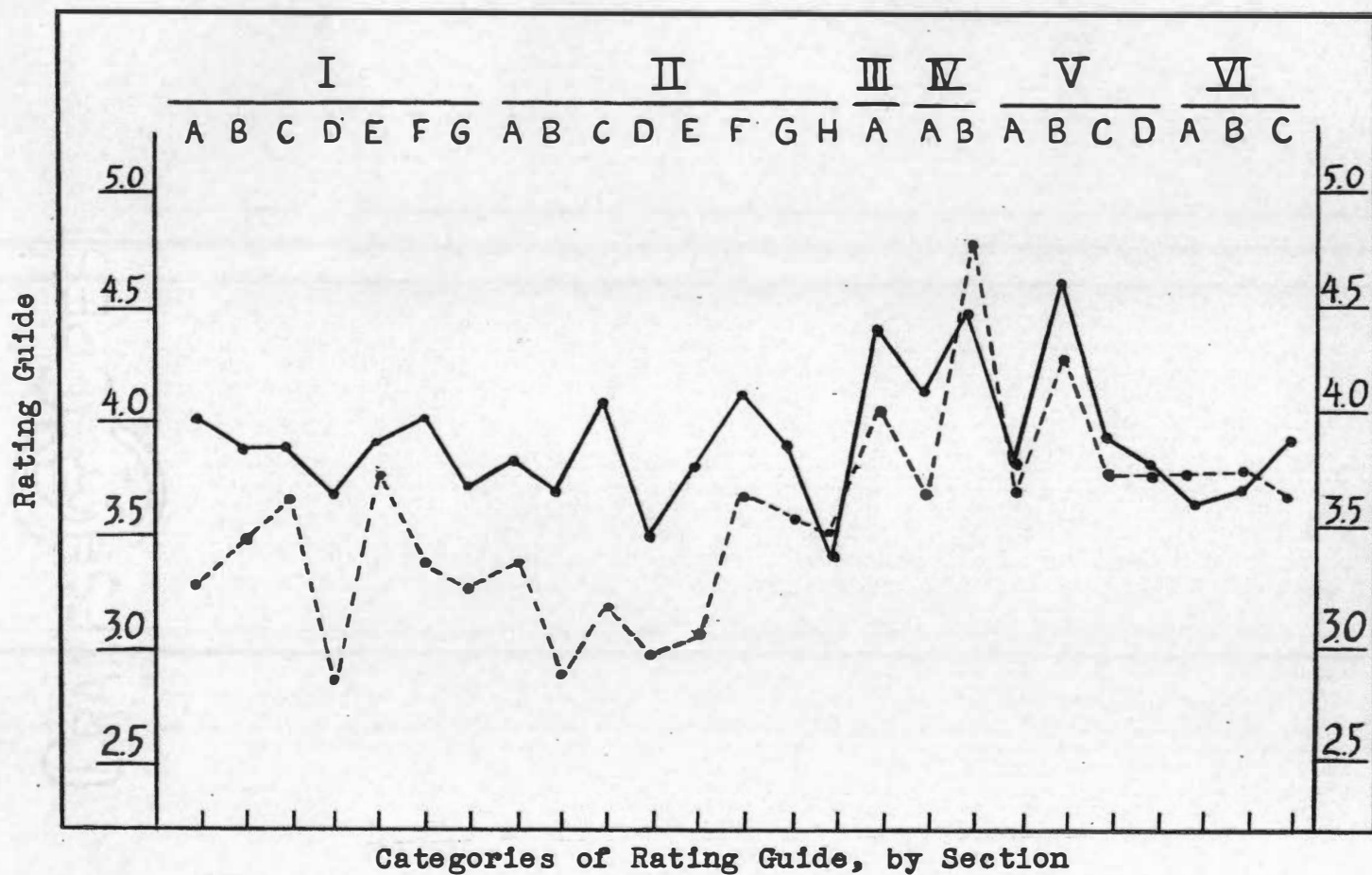


Figure 3

Averages for 1957 Group by Rating Guide Category, Showing
1957 Pre-rating and 1958 Post-rating

—— Post-rating
----- Pre-rating

TABLE IV

PRE-RATING AND POST-RATING AVERAGES FOR 1957 GROUP
BY RATING GUIDE CATEGORIES

Category	Pre-rating	Post-rating	Difference
I-A	3.3	4.0	0.7
I-B	3.5	3.9	0.4
I-C	3.7	3.9	0.2
I-D	3.8	3.9	0.1
I-E	3.4	4.0	0.6
I-F	3.3	3.7	0.4
II-A	3.4	3.8	0.4
II-B	2.9	3.7	0.8
II-C	3.2	4.1	0.9
II-D	3.0	3.5	0.5
II-E	3.1	3.8	0.7
II-F	3.7	4.1	0.4
II-G	3.6	3.9	0.3
II-H	3.5	3.4	- 0.1
III-A	4.1	4.4	0.3
IV-A	3.7	4.1	0.4
IV-B	4.8	4.5	- 0.3
V-A	3.7	3.8	0.1
V-B	4.3	4.6	0.3
V-C	3.8	3.9	0.1
V-D	3.8	3.8	0.0
VI-A	3.8	3.6	- 0.2
VI-B	3.8	3.7	- 0.1
VI-C	3.7	3.9	0.2
Total difference			7.1
Mean difference			0.284*

*Mean difference was not significant at the 0.05 level.

by means of the t-test for paired differences and were found to be non-significant at the 0.05 level.

Section I, Interpersonal Relations

How does he relate to others? Change from an average of 3.3 to 4.0 indicates change from a reserved disposition to one that would be willing to initiate friendships. The first rating indicates a person dependent upon others for the beginning of an acquaintanceship; the latter rating indicates an awareness of other people.

Does he utilize the opinion of others? A pre-rating average of 3.5 indicates an independent sort of existence, a willingness to behave on the basis of one's own judgment. Such a person would accept good suggestions from others but would not solicit them. A post-rating of 3.9 suggests a curiosity about opinions of others, but a high degree of selectivity in the use of other people's ideas.

Is he skillful in developing an organization in which each can do his best? A change from 3.7 to 3.9 is so small that the indication would be that such a person would continue to be fairly subjective about the delegation of responsibilities. Sometimes he would give attention to special interests and abilities of associates.

Is he skillful in continuous implementation of policies? Both the pre-rating of 3.8 and the post-rating of 3.9 indicate a fairly high degree of skill in this category.

This person is somewhat cautious but does move forward in the implementation of policies.

Does he help the group arrive at a working consensus?

Both the pre-rating of 3.4 and the post-rating of 4.0 indicate a desire to have group approval of action. The first rating suggests a tendency to rush action at times without giving enough time for proper group consideration; whereas, the second rating indicates a willingness to allow a group to have time to reach a decision but perhaps indicates offering a minimum of help in procuring the data which would furnish the basis for a reasonable decision.

Does he believe that democratic processes are essential? An indication of growth is present in the change from 3.3 to 3.7. The latter person would like to use democratic processes, but is unable to identify them. His failures are due to lack of knowledge and skill rather than to autocratic intentions. The person of the first rating uses democratic processes when they do not interfere with his normal method of operation.

Section II, Intelligent Operation

Does he give sufficient consideration to new data in problem solving? A tendency to improve is indicated by the change from 3.4 to 3.8 in this category. The first rating indicates a rather casual interest in changes that are taking place in the environment. A 3.8 rating suggests that the

person has developed interest in some areas and that he watches them with a rather consistent concern.

Does he recognize and define problems? A person with either the pre-rating of 2.9 or the post-rating of 3.7 knows that problems exist. The first rating indicates a rather generalized state of confusion. Such a person calls symptoms and problems by the same name. In the second state he can recognize problems as problems, but he still has difficulty in understanding possible causes or proposing possible solutions.

Is he consistent in terms of his basic assumptions? The person with a pre-rating of 3.2 may be fearful of unpopularity or he may not know his own mind, but he frequently displays uncertainty on controversial subjects. Something has happened to give him more stamina by the time of the post rating because the post-rating of 4.1 indicates that the areas he considers important merit an attitude of certainty and consistency on his part. He may still vacillate in areas he considers unimportant.

Does he experiment? Both ratings indicate a rather conservative individual. The first rating of 3.0 indicates a somewhat spasmodic interest in undertaking new projects. This person may be spurred into competitive action when he observes that somebody else has undertaken a new project. The post-rating of 3.5 possibly indicates a little more activity

but without much planning or weighing of significance.

Does he try to recognize and deal with his own biases?

A change from 3.1 to 3.8 indicates strengthening of the quality of self-examination. The pre-rating suggests a person who will admit he is wrong, if he is pushed to do so; he is not always comfortable about his position. The post-rating of 3.8 indicates a person who generally evaluates his position; however, under pressure he is likely to resort to biases.

Does he appear to have profited by previous mistakes?

The change in a pre-rating of 3.7 to a post-rating of 4.1 does not indicate much more than a possible trend. This person may repeat a mistake but generally improves as a result of past mistakes.

Does he have the ability to size up people? Rather traditional behavior is indicated by both the pre-rating of 3.6 and the post-rating of 3.9. This person counts on the knowledge he has gained about people in previous personal experiences to help him form judgments about all people. Whenever problem situations arise, he does seek additional resources for making judgments.

Does he accept responsibility wisely? The person of the pre-rating of 3.5 and the post-rating of 3.4 seems to be a stereotype of a "school person." He concentrates on school routine. There are a few non-school endeavors that he chooses to support, on a highly selective basis.

Section III, Emotional Stability

Is he emotionally stable? The fictional person who is representing the experimental group as a whole is a rather stable individual. He was like that with a pre-rating of 4.1 and more so with a post-rating of 4.4. He meets novel situations calmly. He is developing the ability to help others feel at ease in his presence. He can be distracted emotionally by occasional problems.

Section IV, Ethical and Moral Strength

Does he have the courage of his convictions? The change from a pre-rating of 3.7 to a rating of 4.1 just suggests a trend. Generally this person has well-tempered convictions and tries to follow them. Sometimes he operates with a little uncertainty if he is not convinced of the soundness of some of his convictions.

Does he exhibit integrity in dealing with others? A pre-rating of 4.8 and a post-rating of 4.5 indicate a rather dependable person. Most agreements are regarded as important by him; he may be a little careless about lesser commitments.

Section V, Adequacy of Communication

How well does he express himself orally? The person with a first rating of 3.7 and a second rating of 3.8 has a fairly practical, every-day speech. Either he does not care to or he can not express abstractions well. He is no

linguist; he uses a rather limited vocabulary.

Is he a good listener? Even though the fictional character, the average person in the group, is not impressive as a speaker, he is an exceptionally good listener. His pre-rating of 4.3 and his post-rating of 4.6 indicate an attentive listener. Not only does he listen to expressions of ideas about subjects in which he is interested, but he seems desirous of broadening his scope by listening to ideas about which others have concern.

Does he interest people in examining ideas? This person again indicates a "practical" nature. In the pre-rating of 3.8 and post-rating of 3.9 there is an indication that this person may be some kind of "pep leader" who attempts to rally interest in causes or ideas he personally considers important.

How skillfully does he lead group discussion? This person lacks skill in drawing out representative ideas from members of a group for group consideration. He does not know how to bring about a climate that is conducive to good group work. With a rating of 3.8 both times he indicates the need for a structured agenda, which he can manage very well.

Section VI, Operation as a Citizen

Does he help people interpret significant contemporary trends and events? A pre-rating of 3.8 and a post-rating of 3.6 suggest a person of limited interest in current events; therefore such a person would not be likely to attempt to

encourage others' interest in contemporary affairs. These ratings suggest that the person may know about current affairs but he fails to discuss them.

Is he cooperative with non-educational groups working for community betterment? A pre-rating of 3.8 and a post-rating of 3.7 indicate a person who is probably aware of the need for a "good neighbor policy." But his efforts to cooperate with community groups may be spasmodic or may result in his spreading his efforts so thinly they are ineffective. He is conscious of his need to cooperate but may use poor judgment as to how to do it.

What is his attitude toward minority groups in the school community? This person is conscious that the public school is obligated to serve the children in the entire community. He prefers to avoid personal involvement in any controversial issues concerning minority groups. The pre-rating of 3.7 and the post-rating of 3.9 suggest that this person will uphold the right of most minority viewpoints to be represented, but he may be reluctant to recognize those that are considered extreme.

Group Change by Rating Guide Sections

Table V lists the pre-rating and post-rating average scores of the 1957 experimental group on the various sections of the Rating Guide. As the table indicates, the group seemed

TABLE V

THE AVERAGE SCORES OF THE 1957 EXPERIMENTAL GROUP ON THE
VARIOUS SECTIONS OF THE RATING GUIDE AS INDICATED
BY PRE-RATINGS AND POST-RATINGS

Section	Average Scores		
	Pre- rating	Post- rating	Differ- ence
I. Interpersonal Relations	3.4	3.9	0.5
II. Intelligent Operation	3.3	3.8	0.5
III. Emotional Stability	4.1	4.4	0.3
IV. Ethical and Moral Strength	4.3	4.3	0.0
V. Adequacy of Communication	4.0	3.9	- 0.1
VI. Operation as a Citizen	3.8	3.7	- 0.1
Total difference			1.1
Mean difference			0.183*

*Mean difference was not significant at 0.05 level.

to change most with respect to Interpersonal Relations, Intelligent Operation, and Emotional Stability. Little change was shown with respect to Ethical and Moral Strength, Adequacy of Communication, and Operation as a Citizen.

Differences between the pre-rating and post-rating scores were obtained. An average change of 0.183 units and a variance of 0.0806 were obtained. These differences were analyzed by means of the t-test for paired differences and found to be non-significant at the 0.05 level.

The composition of the group in relation to change between ratings by Rating Guide Sections is shown in Table VI. a description of the group change by Sections follows.

Section I, Interpersonal Relations

The group rating changed from 3.4 to 3.9. An increased awareness of the importance of recognizing the abilities and the worth of other individuals is suggested. This change indicates more willingness to allow participation in decision-making of those involved. Thirteen persons showed a positive change in this area; three showed a negative change; and two indicated no change.

Section II, Intelligent Operation

The group changed from 3.3 to 3.8 on this section of the Rating Guide. A trend toward the problem-solving approach to functioning is suggested. An increased willingness to base

TABLE VI

NUMBERS AND PERCENTAGES SHOWING CHANGE BY SECTIONS OF THE
RATING GUIDE AS INDICATED BY PRE-RATINGS AND POST-
RATINGS OF THE 1957 EXPERIMENTAL GROUP

Rating Guide Section	Number Who Changed			Percentage Who Changed		
	Posi- tive	No Change	Nega- tive	Posi- tive	No Change	Nega- tive
Section I	13	2	3	72	11	17
Section II	13	0	5	72	0	28
Section III	8	7	3	44	39	17
Section IV	7	5	6	39	28	33
Section V	10	2	6	56	11	33
Section VI	7	3	8	39	17	44

decisions on available data is indicated. Thirteen persons showed change in a positive direction, while five showed a negative change.

Section III, Emotional Stability

This section of the Rating Guide has but one category; therefore, the changes noted here are the same as those noted for the category earlier in this chapter. The change from 4.1 to 4.4 is probably significant because the latter rating is approaching the upper limit of 5.0 on the Rating Guide. The rating of 4.4 indicates the ability to exhibit calmness in a helpful manner. Eight persons showed an increased rating; seven showed no change (all these, however, were at the 4.0 or the 5.0 level); and three showed a negative change.

Section IV, Ethical and Moral Strength

This is a two-category section. An average may not be descriptive. In the first category a positive change occurred; in the second category a negative change occurred. The first was an indication that the group had increased its courage about its convictions; the second was an indication that there was increased carelessness in regarding agreements with others as a matter of integrity. The pre-rating average and the post-rating average were the same, 4.3. Seven persons improved in the total area of ethical and moral strength; six persons showed a decreased score; and five persons remained the same.

Section V, Adequacy of Communication

The pre-rating of 4.0 and the post-rating of 3.9 indicate that the group is fairly well skilled in a practical, day-to-day method of communicating. A conscious effort to increase professional responsibility in the area of communication is not indicated. Ten people showed increased ability in this area; six showed decreased ability; and two did not change.

Section VI, Operation as a Citizen

The pre-rating of 3.8 and the post-rating of 3.7 indicate little more than average community leadership. Some "intellectual" interest in the affairs of citizenship is shown, but in practice there is evidence of little activity. Seven persons showed positive change; eight persons showed negative change; and three seemed not to change.

A Comparison of Change in This Group with Change
in an Earlier Group as Indicated by
the Tennessee Rating Guide

General Comparison of Groups

No conspicuous differences in the 1956 pilot group and the 1957 experimental group existed in the areas of age, experience, or positions. The 1956 pilot group of seventeen students was composed entirely of men; the 1957 group of eighteen students (who participated throughout the study)

included six women. The pre-rating average of the 1956 group was 3.4; the pre-rating average of the 1957 group was 3.6. The post-rating average of the 1956 group was 3.6; the post-rating average of the 1957 group was 3.9. The rating averages were based on the total Rating Guide. As indicated above, greater change occurred in some sections of the Rating Guide than in others in the 1957 group. The same thing was true of the 1956 group.

Both groups seemed to change in a positive direction with respect to characteristics involving Interpersonal Relations, Intelligent Operation, and Emotional Stability. The 1956 group showed positive change also in respect to Operation as a Citizen and Adequacy of Communication. Some changes were made in the Rating Guide between the 1956 and the 1957 ratings. Adjustments based on the Rating Guide modifications were made in the treatment of data for comparison purposes. The total number of categories in the Rating Guide was twenty-eight when the 1956 ratings were made; the total number was twenty five when the 1957 ratings were made.

Comparison by Categories

Of the categories that ranked in the top seven (of twenty-five) in relation to total amount of positive change, four were common to both groups. These were as follows: Section I, Category F, Does he help the group arrive at a working consensus? Section II, Category C, Is he consistent

in terms of his basic assumptions, Section II, Category D, Does he experiment? Section II, Category E, Does he try to recognize and deal with his own biases?

The categories that ranked in the top seven in relation to positive change that were peculiar to the 1956 group were: Section I, Category G, Does he believe that democratic processes are essential? Section V, Category A, How well does he express himself orally? Section VI, Category A, Does he help people interpret significant contemporary trends and events? The categories that were in the top seven that were peculiar to the 1957 group were: Section I, Category A, How does he relate to others? Section I, Category D, Is he skillful in getting policies formulated cooperatively?

Categories in which there was either negative change or no change in the 1956 group were Section IV, Category B, Does he exhibit integrity in dealing with others? and Section V, Category B, Is he a good listener? In the 1957 group the categories in which there was negative or no change were: Section II, Category H, Does he accept responsibility wisely? Section IV, Category B, Does he exhibit integrity in dealing with others? Section V, Category D, How skillfully does he lead group discussion? Section VI, Category A, Does he help people interpret significant contemporary trends and events? Section VI, Category B, Is he cooperative with non-educational groups working for community betterment? Category B, Section IV was common to both groups.

Figure 4, page 63, illustrates the comparative change of the two groups by categories.¹ A zero line indicates "no change." Bars drawn away from the zero line in opposing directions indicate changes in positive or negative directions.

Categories with "Effective" Rating

In his study of patterns of behavioral characteristics exhibited by school administrators Gentry concluded that "a rating of four or better characterized the effective administrator" in the group who composed his study.² An examination of both the 1956 group and the 1957 group indicated that each group rated as high as four in three categories in pre-ratings. The 1956 group rated as high as four in four categories in the post-rating; the 1957 group rated as high as four in eight categories in the post-rating.

In pre-ratings two categories were common to both groups: Category B, Section IV, Does he exhibit integrity in dealing with others? and Category B, Section V, Is he a good

¹Gem Kate Taylor, "The Development of Effective Characteristics in Students of School Administration" (Unpublished Ed. D. thesis, The University of Tennessee, June 1957), pp. 82-98.

²Harold Wayne Gentry, "Patterns of Behavioral Characteristics Exhibited by School Administrators" (Unpublished Ed. D. thesis, The University of Tennessee, March 1957).

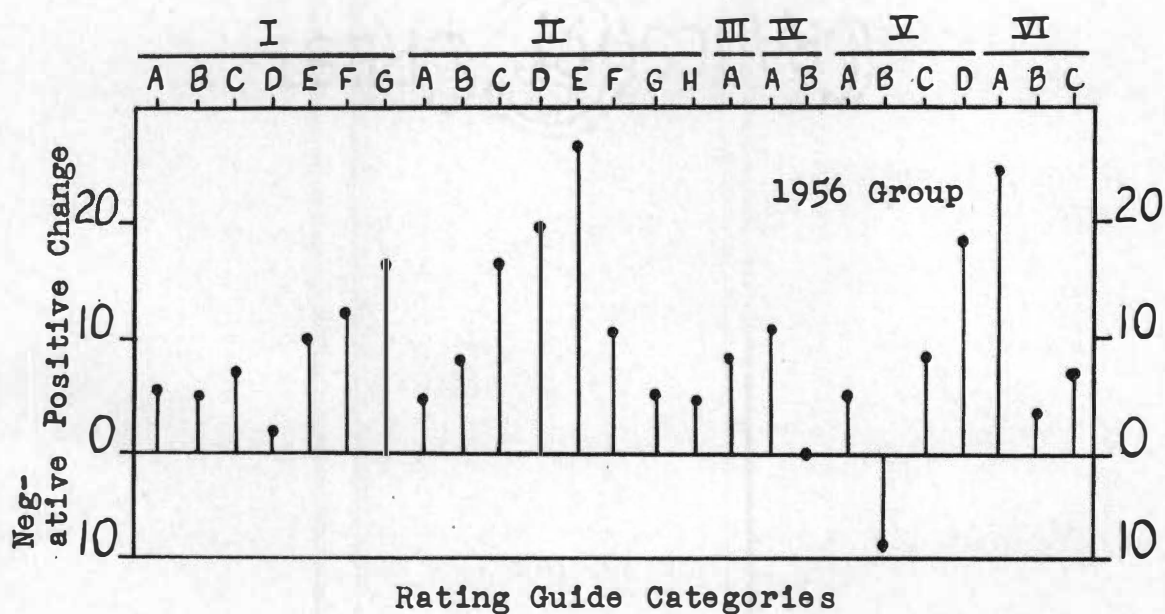
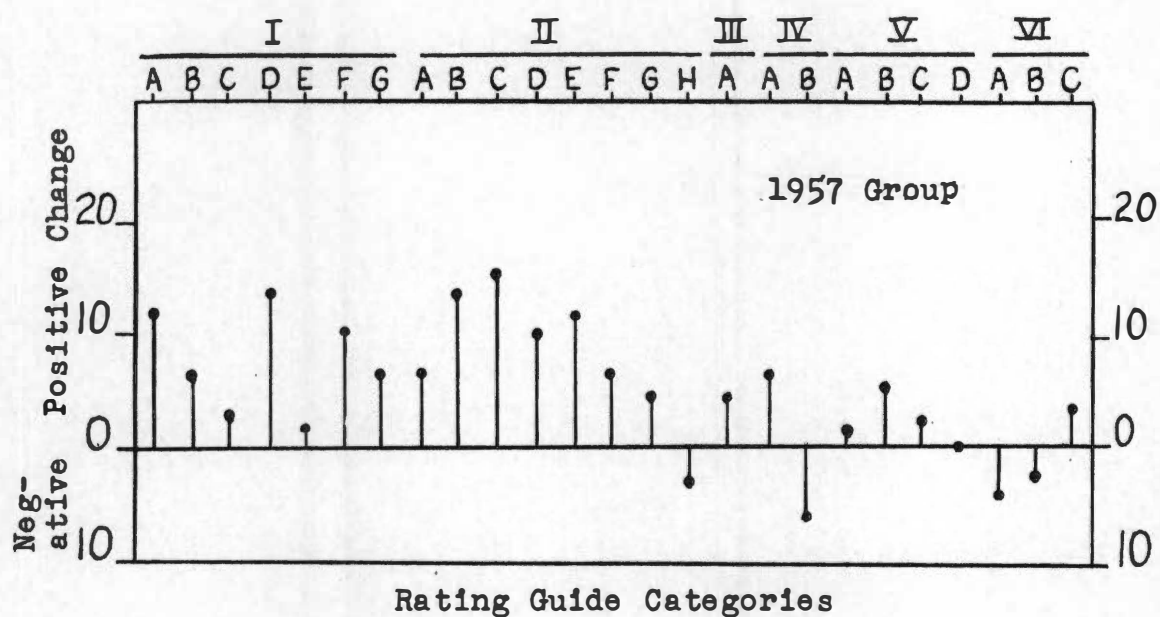


Figure 4

Group Change Between First and Second Year
Ratings for Each Rating Guide Category
for Both the 1956 and the 1957 Groups

listener? The 1956 group had a pre-rating as high as four in Category A, Section IV, Does he have the courage of his convictions? The 1957 group had a pre-rating as high as four in Category A, Section III, Is he emotionally stable?

In the as-high-as-four post-rating of both groups were Category B, Section IV; Category B, Section V; Category A, Section III (the three are described above); and Category F, Section II, Does he appear to have profited by previous experiences? In addition to these categories, the 1957 group rated as high as four in the following categories: Category A, Section IV (described above); Section I, Category A, How does he relate to others? Section I, Category F, Does he help the group arrive at a working consensus? and Section II, Category C, Is he consistent in terms of his basic assumptions?

Chapter Summary

Rating procedures were given in this chapter. Group change by categories and by sections of the Rating Guide were indicated. The chapter concluded with a comparison of change in the 1957 experimental group with an earlier similar experimental group.

CHAPTER IV

DESCRIPTION OF THE GROUP AS INDICATED BY TESTS, AND DIFFERENCES BETWEEN INDIVIDUALS WHO CHANGED AND THOSE WHO DID NOT CHANGE

As stated earlier, certain tests were given at the beginning of the 1957 experimental program for two reasons:

(1) to contribute to self-understanding of the person involved in the experiment and (2) to analyze results for differences between individuals who changed and those who did not change during the period of the study. Neither the changes noted in behavioral characteristics as indicated by the Rating Guide nor the descriptions of individuals as indicated by various test scores were regarded as absolute, but the data were regarded as guides to interpretation about possible change through education.

"Measurement always takes place in a more or less complex situation in which an innumerable variety of factors may affect both the characteristic which is being measured and the process of measurement itself," observes Jahoda.¹ Jahoda summarizes some of the possible sources of differences in scores among a group of individuals as being: (1) true

¹Marie Jahoda, Morton Deutsch and Stuart W. Cook, Research Methods in Social Relations, Part One: Basic Processes (New York: The Dryden Press, 1951), p. 95.

differences in the enduring characteristic which one is attempting to measure; (2) true differences in other enduring characteristics of the individual which affect his score; (3) difference due to (a) transient personal factors, (b) situational factors, (c) inadequate sampling of items, (d) lack of clarity of the measuring instrument, (e) variations in administration, (f) mechanical factors and (g) factors in the analysis.²

How can individuals be identified who would be unlikely to respond to an educational program similar to the experimental programs? Leads to answers may have been furnished through the simultaneous examination of data pertaining to change in individuals and their scores on certain tests. The answers are important in relation to the problem of the selection of graduate students in educational administration.

Test Results

A description of the tests given and the results (on a group basis) are presented in the following pages. The mean scores and percentiles of the pilot group of 1956 are shown. The scores and percentiles of the 1956 group are from Taylor's study.³ Also included are normative means and percentiles,

²Ibid., pp. 95-100.

³Gem Kate Taylor, "The Development of Effective Characteristics in Students of School Administration" (Unpublished

if available. The 1956 group was all men, except for one doctoral student who did not participate as a subject. The 1957 experimental group was composed of both men and women. Some test manuals used in the study furnished separate percentiles for men and women. In such cases both are given; also, in such cases, a total for the group is given.

Allport-Vernon Study of Values

The Allport-Vernon Study of Values purports to measure the relative prominence of six basic motives in personality: theoretical, economic, aesthetic, social, political, and religious.⁴ The scale was designed for use with college students, or with adults who have had some college or its equivalent. The only mean score for the 1957 group that fell outside the "average" limits was the score of 33.7 for aesthetic, which was "low"; however, the lower limit for "average" is 34.0. The range of scores and the mean scores for the 1957 group, the mean scores for the 1956 group and the normative means are given in Table VII.

Ed. D. thesis, The University of Tennessee, June 1957).

⁴Gordon W. Allport, Philip E. Vernon and Gardner Lindzey, Study of Values (Boston: Houghton Mifflin Company, 1951), Manual of Directions.

TABLE VII

MEAN SCORES OF THE 1956 AND 1957 GROUPS
FOR THE ALLPORT-VERNON STUDY OF VALUES

Values	1957 Group		1956 Group Mean	Normative Mean
	Range of Scores	Mean		
Theoretical	24-51	41.7	40.6	44.3
Economic	29-52	40.2	41.2	37.1
Aesthetic	17-54	33.7	33.6	43.8
Social	29-54	39.8	42.8	42.0
Political	27-58	40.8	40.6	39.2
Religious	29-60	43.8	41.6	33.6

Edwards Personal Preference Schedule

The P P S was designed primarily as an instrument for research and counseling purposes, to provide quick and convenient measures of a number of relatively independent normal personality variables. The variables were named according to "manifest needs."⁵ The fifteen personality variables which were measured by the P P S are: Achievement, Deference, Order, Exhibition, Autonomy, Affiliation, Intraception, Succorance, Dominance, Abasement, Nurturance, Change, Endurance, Heterosexuality, and Aggression. The only mean score outside the "average" range was a "low" percentile rating of 14 ("low" range, 4-16) on heterosexuality by the men. This contrasts with a percentile rating of 45 by the men who composed the 1956 group. Perhaps the fact that the 1957 group was of heterosexual composition affected the scores.

Table VIII shows range of scores, mean scores, percentiles for the men, percentiles for the women, and percentiles for the total group for the 1957 group, mean scores, and percentiles for the 1956 group. Although all scores except one were "average," the top four and bottom five percentile ratings are listed here since they may be clues to the composition of the group, personality-wise. The top

⁵Allen L. Edwards, Edwards Personal Preference Schedule (New York: The Psychological Corporation, 1954).
Manual of Directions.

TABLE VIII

MEAN SCORES OF THE 1956 AND 1957 GROUPS FOR
THE EDWARDS PERSONAL PREFERENCE SCHEDULE

Needs	1957 Group					1956 Group	
	Range of Scores	Mean Scores	Percentiles Men	Percentiles Women	Percentiles Total	Mean Scores	Percentiles
achievement	10-25	15.3	50	79	60	11.8	16
deference	9-20	14.3	81	80	81	14.4	88
order	5-18	11.7	71	64	69	11.9	63
exhibition	6-20	12.3	37	21	32	13.0	37
autonomy	4-20	10.4	28	28	28	23.1	43
affiliation	7-25	16.2	72	41	62	16.2	64
intraception	9-26	19.0	64	80	69	16.2	51
succorance	3-22	10.7	50	51	50	12.1	65
dominance	9-25	16.8	45	67	52	13.1	32
abasement	3-23	13.0	53	74	60	12.9	53
nurturance	8-25	17.6	82	74	79	13.3	46
change	6-24	14.5	41	35	39	14.9	41
endurance	5-24	14.6	69	57	65	13.5	57
heterosexuality	2-26	11.0	14	24	17	17.4	45
aggression	2-22	12.5	57	51	55	12.5	47

Percentiles

97 and above
85-96
17-84
4-16
3 and below

Interpretation

Very high
High
Average
Low
Very low

four in descending order were nurturance for men, deference for men, deference for women and intraception for women. The percentiles ranged from 82-80 for these variables. The bottom five in ascending order were heterosexuality for men, exhibition for women, heterosexuality for women, and autonomy for both men and women. The percentiles ranged from 14-28 for these variables.

Partial definitions of the needs associated with the variables mentioned are:

Nurturance: to help friends when they are in trouble, to show a great deal of affection toward others, to have others confide in one about personal problems.

Deference: to accept the leadership of others, to conform to custom and avoid the unconventional, to find out what others think.

Intraception: to put one's self in another's place to judge people by why they do things rather than by what they do, to analyze one's motives and feelings.

Heterosexuality: to go out with members of the opposite sex, to engage in social activities with the opposite sex, to be in love, to be regarded as physically attractive by members of the opposite sex.

Exhibition: to be the center of attention; to say things just to see what effect it will have on others, to tell amusing jokes and stories.

Autonomy: to be independent of others in making decisions, to avoid situations where one is expected to conform, to be able to come and go as desired.

Thurstone Temperament Schedule

The Thurstone Temperament Schedule was devised for the purpose of measuring important stable traits which describe how normal, well-adjusted people differ from each other.⁶ It was limited to a description of important aspects of temperament and was not designed to appraise the degree of conflict, insecurity or maladjustment. Seven areas of temperament were appraised in a relatively short questionnaire. The areas were: Active, Vigorous, Impulsive, Dominant, Stable, Sociable, and Reflective.

Table IX shows range of scores, mean scores, percentiles for the men, percentiles for the women and percentiles for the total for the 1957 group, the mean scores, and the percentiles for the 1956 group. The only "low" and the only "high" ratings were made by the women. The "low" was in the impulsive area. The "high" was in the stable area. The group as a total did not rate below the fiftieth percentile except in the impulsive area.

⁶L. L. Thurstone, Thurstone Temperament Schedule (Chicago: Science Research Associates, Inc., 1953). Manual of Directions.

TABLE IX

MEAN SCORES OF THE 1956 AND 1957 GROUPS
FOR THE THURSTONE TEMPERAMENT SCHEDULE

Traits	Range of Scores	1957 Group				1956 Group	
		Mean Scores	Percentiles Men	Percentiles Women	Percentiles Total	Mean Scores	Percentiles
Active	4-16	10.4	60	51	57	10.5	47
Vigorous	6-16	10.8	51	72	58	11.5	40
Impulsive	4-15	9.7	30	20	27	8.7	13
Dominant	0-20	11.1	61	71	64	10.0	53
Stable	6-19	12.4	66	82	71	11.0	57
Sociable	7-19	13.1	52	62	55	12.0	42
Reflective	3-15	8.9	56	61	58	8.0	43

A similarity between the 1956 and the 1957 groups occurs in the Thurstone Temperament Schedule measurements. The extreme scores were in the same areas for both groups: they scored lowest in the impulsive area and the highest in the stable area. A low score on impulsive indicates a person who doggedly hangs on" when acting or thinking. This person is rather slow and deliberate when making decisions. Persons who have high stable scores usually are cheerful and have even dispositions. They remain calm in a crisis; they do not fret about daily chores; they are not annoyed by leaving a task unfinished or by having to finish it by a deadline.

Other Tests

Cooperative English. Two parts of the Cooperative English test were given: Part A, Mechanics, which measured the degree of mastery of the essentials of correct English usage; and Part B, Effectiveness of Expression, which measured the skills of sentence structure and style, active vocabulary, and organization.⁷ The mean for the mechanics section was 57.0; the mean for the effectiveness of expression section was also 57.0. A mean of 57.1 for mechanics had been obtained for seniors in colleges that were representative of "many junior and teachers" colleges; a mean of 58.5 was

⁷Cooperative English Test, Percentile Ranks for High School and College Students (New York: The Cooperative Test Service, 1940). Manual

obtained for the same group of college seniors in the effectiveness of expression section. The 1956 group had means of 38.0 and 52.8 for the respective sections.

Minnesota Teacher Attitude Inventory. The Minnesota Teacher Attitude Inventory was designed to measure those attitudes of a teacher which will predict how well he is likely to get along with pupils in interpersonal relationships, and, indirectly, how well satisfied he will be with teaching as a vocation.⁸ The teacher ranking at the high end of the scale is predicted to be able to maintain a classroom climate which would favor democratic behavior. At the other extreme of the scale would be the prediction that the teacher would attempt to dominate the classroom situation. This test measures the teacher's attitude toward democratic operation.

The test was administered at the beginning of the summer term and at the end of the term. The correct interpretation of the change in scores would be difficult to make. A great deal of verbal change took place. The mean score of the first test was 45.8; the mean score of the second test was 96.8. To calculate a normative mean would have been a hazardous task, involving numerous factors such as "years of experience," whether experience was "rural or urban," and "the size of the school" for each member of the group. The

⁸Walter W. Cook, Minnesota Teacher Attitude Inventory (New York: The Psychological Corporation, 1951). Manual of Directions.

means of the 1956 group for test and retest were 58.1 and 92.2, respectively.

P M S Scale. The P M S Scale is a questionnaire which was constructed in the Department of Psychology of the University of Tennessee. It measured the degree to which one employs magical thinking. It was based on a theory that people often act in a way that is different from what they believe, on a knowledge basis. Feelings and reactions toward environmental events and other factors which affect one's life are frequently based on superstitions which exist in the culture; the individual may not be consciously aware of the cultural influence. The mean for P (power) was 35.2; the mean for M (magic) was 74.4; and the mean for S (superstition) was 52.9. The standardized mean was 43.0 for P, 83.7 for M and 52.9 for S. The scores for the 1956 group were 47.0 for P, 75.7 for M, and 40.7 for S.

Professional Beliefs Inventory. This inventory was a study of professional beliefs. Two portions of the inventory were administered to the group. The first dealt with Child Growth and Development and the other with The School in a Democratic Society. The mean score for the first section was 53.0 and for the second section was 33.0. No normative means were available. The means for the 1956 group were 58.1 and 31.3. The total scores for the two groups were similar: 86.0 for the 1957 group and 89.4 for the 1956 group.

Watson-Glaser Critical Thinking Appraisal. The Watson-Glaser Critical Thinking Appraisal was a test which measured an individual's capacity for critical thinking. The test was designed with statements of fact from which the individual being tested was required to draw inferences, to reason deductively, to recognize assumptions, to interpret facts, and to evaluate arguments. The test was given at the beginning and at the end of the summer program. Remembering such a test would be difficult, if not impossible; therefore, changes in scores probably meant changes in ability.

Comparisons of the groups were made with percentile norms for college students. The college norms were obtained from the distribution of scores of the freshmen in a large Eastern university who applied for classification as sophomores in 1951. Norms were not presented separately by sex since the evidence indicated that the differences in scores between sexes had no significant distribution. According to a commonly used five-level classification, percentiles of 94 and above may be considered as "Very High," of 70-93 as "High," of 32-69 as "Average," of 8-31 as "Low," and of 1 through 7 as "Very Low."⁹

⁹Goodwin Watson and Edward M. Glaser, Watson-Glaser Critical Thinking Appraisal (Yonkers-on-Hudson: World Book Company, 1952), p. 6. Manual of Directions.

For the 1957 group the mean score for the first test was 72.5, the fifty-ninth percentile for college students, a classification of "Average"; the mean score for the second test was 75.7, the seventy-fifth percentile, a classification of "High." The 1956 group scored as follows: a mean score of 64.7, the twenty-ninth percentile, or "Low" for the first test; a mean score of 71.8, the fifty-sixth percentile, or "Average" for the second test.

Change as Indicated by the Rating Guide

Figure 5, page 79, shows change that occurred within the period of one year. Individual changes in both the 1956 pilot group and the 1957 experimental group are indicated in the figure.

The problem of dealing statistically with ordinal scales was discussed in Chapter II. The difficulty of "building interval scales" in the area of human traits as represented by the Thurstone methods is observed by Jahoda in these words, "There seems to be considerable doubt that the 'equal-appearing' intervals are, in fact, equal intervals."¹⁰ For the purpose of separating students into upper, middle, and lower groups in regard to change, the ordinal

¹⁰Jahoda, op. cit., p. 194.

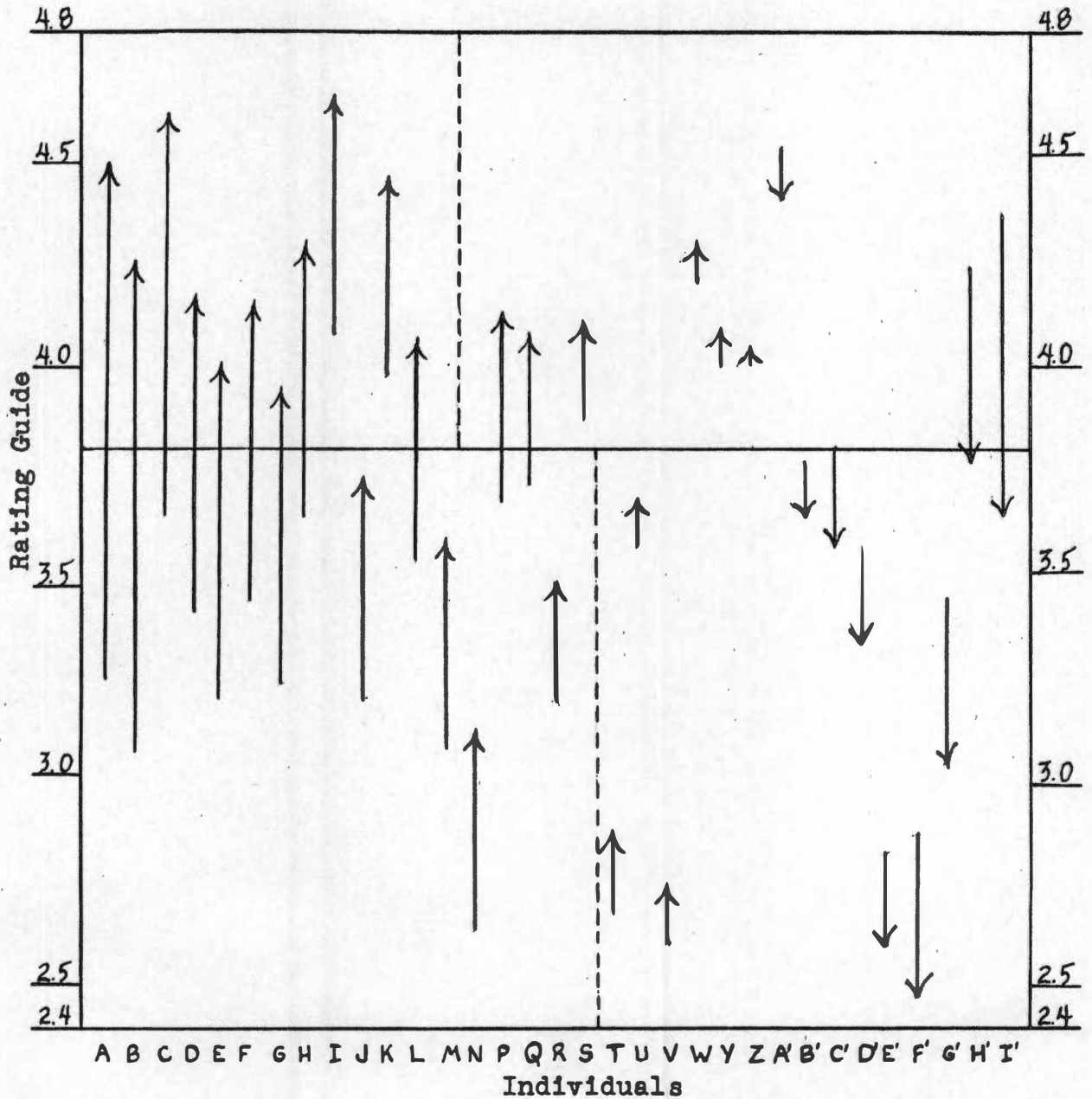


Figure 5

Rating Guide Averages of Individuals in Both the 1956 and the 1957 Groups Showing Change Between Pre-rating and Post-rating for Each Individual, Showing Upper and Lower Thirds of Group

scale of the Rating Guide was assumed to be adequate in the present study.

The upper group was defined as "those whose post-ratings fell above the group post-rating mean and who fell in the upper third in regard to amount of total change, as measured by pre-ratings and post-ratings." The lower group was defined as "those whose post-ratings fell below the group post-rating mean and who fell in the bottom third in regard to amount of total change, as measured by differences in pre-ratings and post-ratings." The bottom third was composed of those who made a negative change, no change, or "little change in relation to the change made by the upper group." The middle group is not defined because no study was made of those individuals, as a separate group.

Correlation of Change and Test Results

A cursory examination of data pertaining to age or number of years of experience showed no relationship to test scores except, to some degree, those of the Minnesota Teacher Attitude Inventory. In the 1957 group four of the five low scores were made by students who had had one year's experience or none. The high test scores were made by principals.

Following the grouping of individuals on the basis of change as indicated by the Rating Guide, the test scores were examined to see if characteristics of the upper group differed

from characteristics of the lower group. As Jahoda points out, one "must constantly be concerned with the social and psychological meaning of his findings as well as their statistical significance."¹¹

Ten tests or parts of tests suggested that their use might be a source for further research. The relationships between those test scores and changes as indicated by the Rating Guide were plotted graphically. The graphs suggested correlations, visually. Results were widely scattered; the interpretation that use of the tests may be a source for further research may be spurious. Figures 6 and 7, pages 82 and 83, show tests that may have a positive correlation to change. The tests that tended to show a positive correlation were: dominance section, Edwards Personal Preference Schedule; Minnesota Teacher Attitude Inventory; Watson-Glasler Critical Thinking Appraisal; active, dominant and stable sections, Thurstone Temperament Schedule.

Figure 8, page 84, shows tests that may have a negative correlation to change. ("Change" when used in this study means "change in the direction of effective behavioral characteristics as identified in the Rating Guide.") The tests that tended to show a negative correlation were: succorance and abasement sections, Edwards Personal Preference Schedule; and P and M sections, P M S Scale.

¹¹Ibid., p. 286.

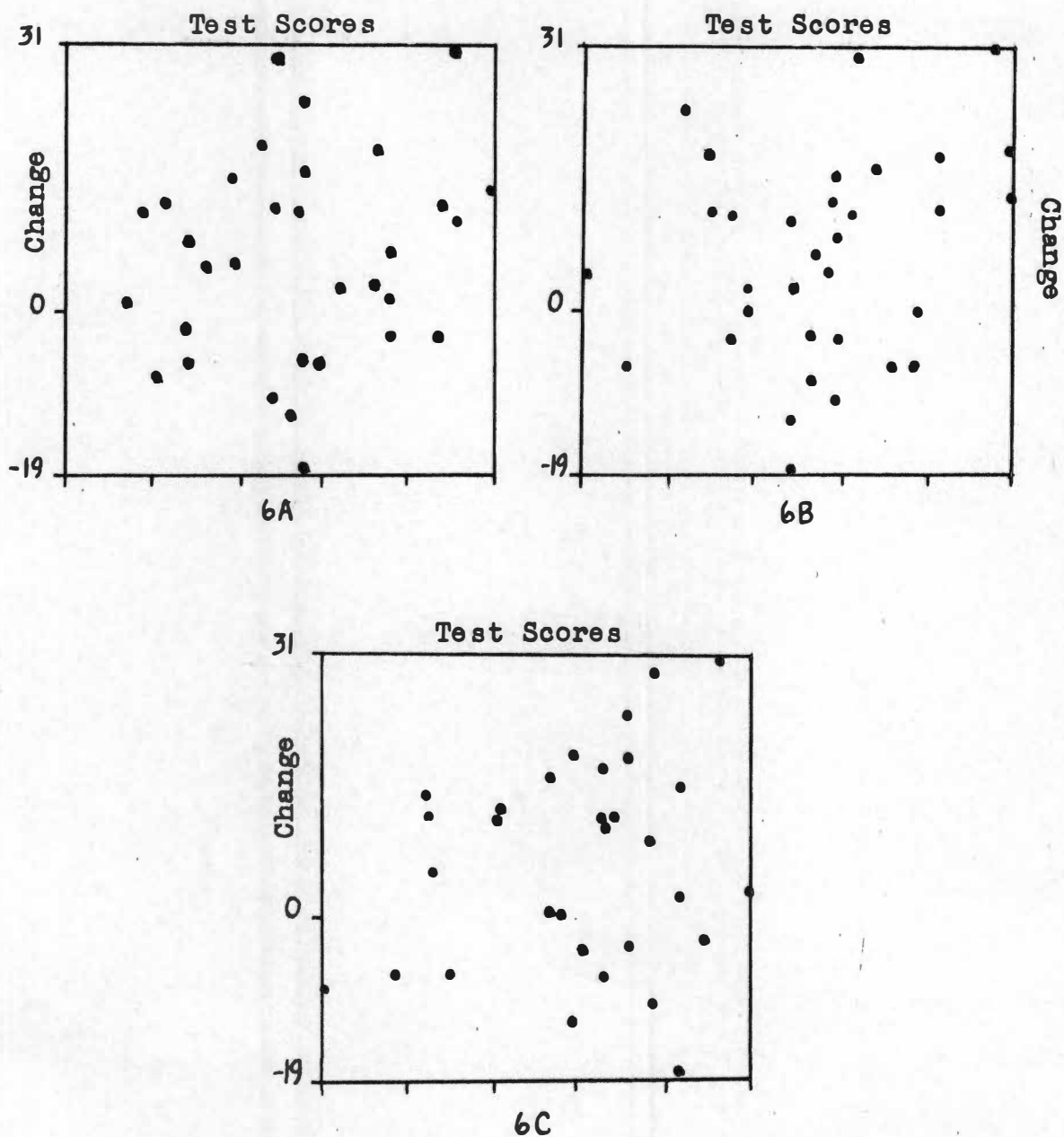


Figure 6

Graphs Suggesting Positive Correlation Between Great Amount of Change Between Pre-rating and Post-rating and High Scores on Certain Tests, a. dominance, Edwards, b. Minnesota Teacher Attitude, c. Watson-Glaser Critical Thinking

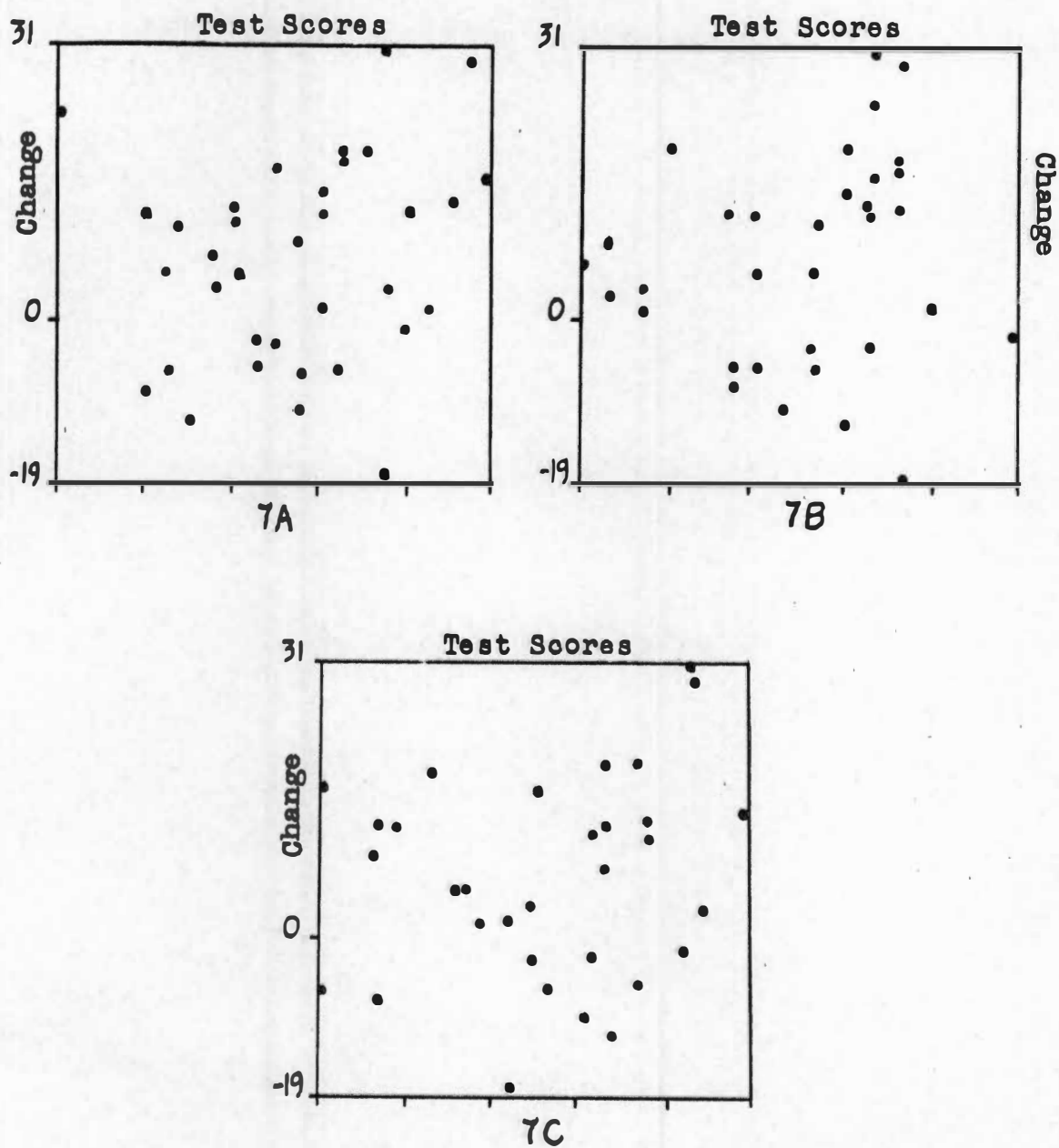


Figure 7

Graphs Suggesting Positive Correlation Between Great Amount of Change Between Pre-rating and Post-rating and High Scores on Certain Tests, a. active, Thurstone, b. dominant, Thurstone, c. stable, Thurstone

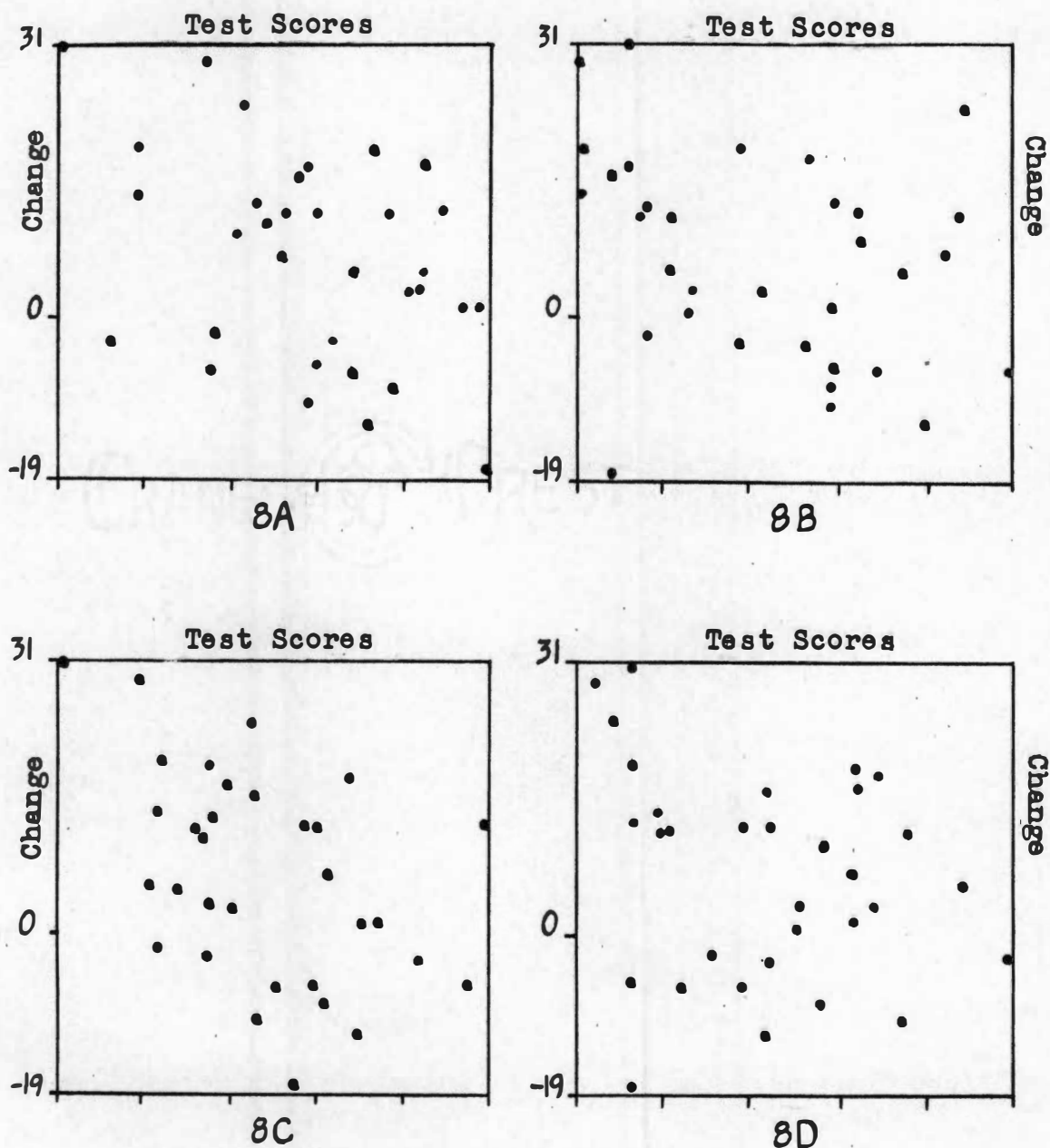


Figure 8

Graphs Suggesting Negative Correlation Between Great Amount of Change Between Pre-rating and Post-rating and High Scores on Certain Tests, a. succorance, Edwards, b. abasement, Edwards, c. P, PMS Scale, d. M, PMS Scale

Important Differences

Differences in Test Score Means

The means of certain tests for the upper group and the lower group were distinctly different. The tests were those listed in the preceding section of this chapter. A scaled chart, Figure 9, page 86, is given that includes ranges of test scores of the total group (the combined 1956 and 1957 groups).

In the first six scales, reading from left to right, the high scores are shown at the top of the figure; in the last four scales the low scores are shown at the top of the figure. In the first six scales high scores indicate those who change; in the last four scales low scores indicate those who change. The means for the upper group of students are shown by a solid line; the means for the lower group are shown by a broken line.

A hypothesis which might bear further examination is that individual profiles based on the scales in Figure 9 might be plotted for the purpose of predicting change in an individual. The individual profile could be placed against the profile for the upper group mean as a guide in predicting ability to change. Four selected profiles from the group (two from the upper group and two from the lower group) are given in Figure 10, page 87. The solid line shows the mean for the upper group in each case; the individual's profile is indicated by the broken line.

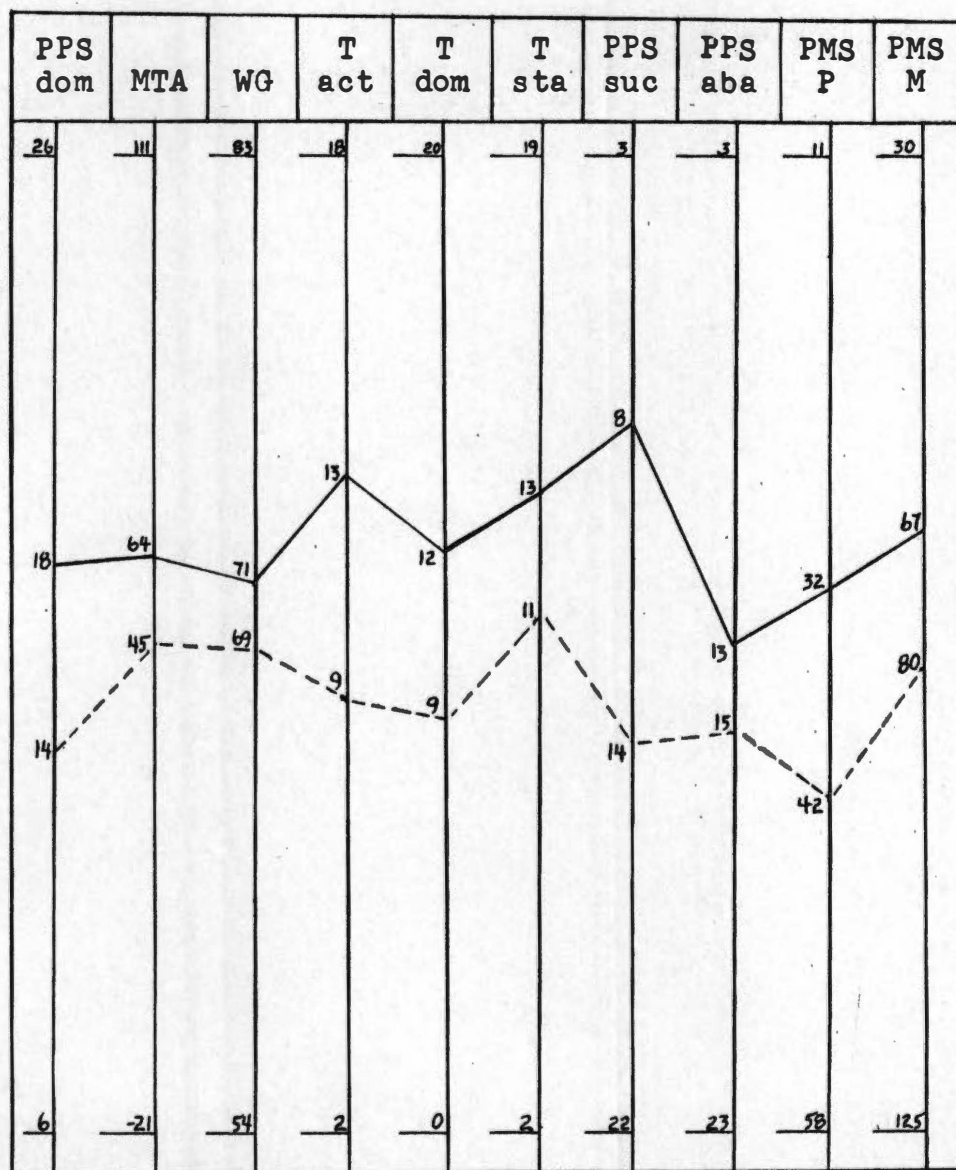


Figure 9

Scale Showing Test Range of Combined 1956 and 1957 Groups for Ten Tests, dominance, Edwards; Minnesota Teacher Attitude; Watson-Glaser Critical Thinking; active, dominant, stable, Thurstone; succorance, abasement, Edwards; P, M, PMS Scale (Mean Score Profiles are Indicated for Upper and Lower Thirds)

—— Upper Group
 - - - Lower Group

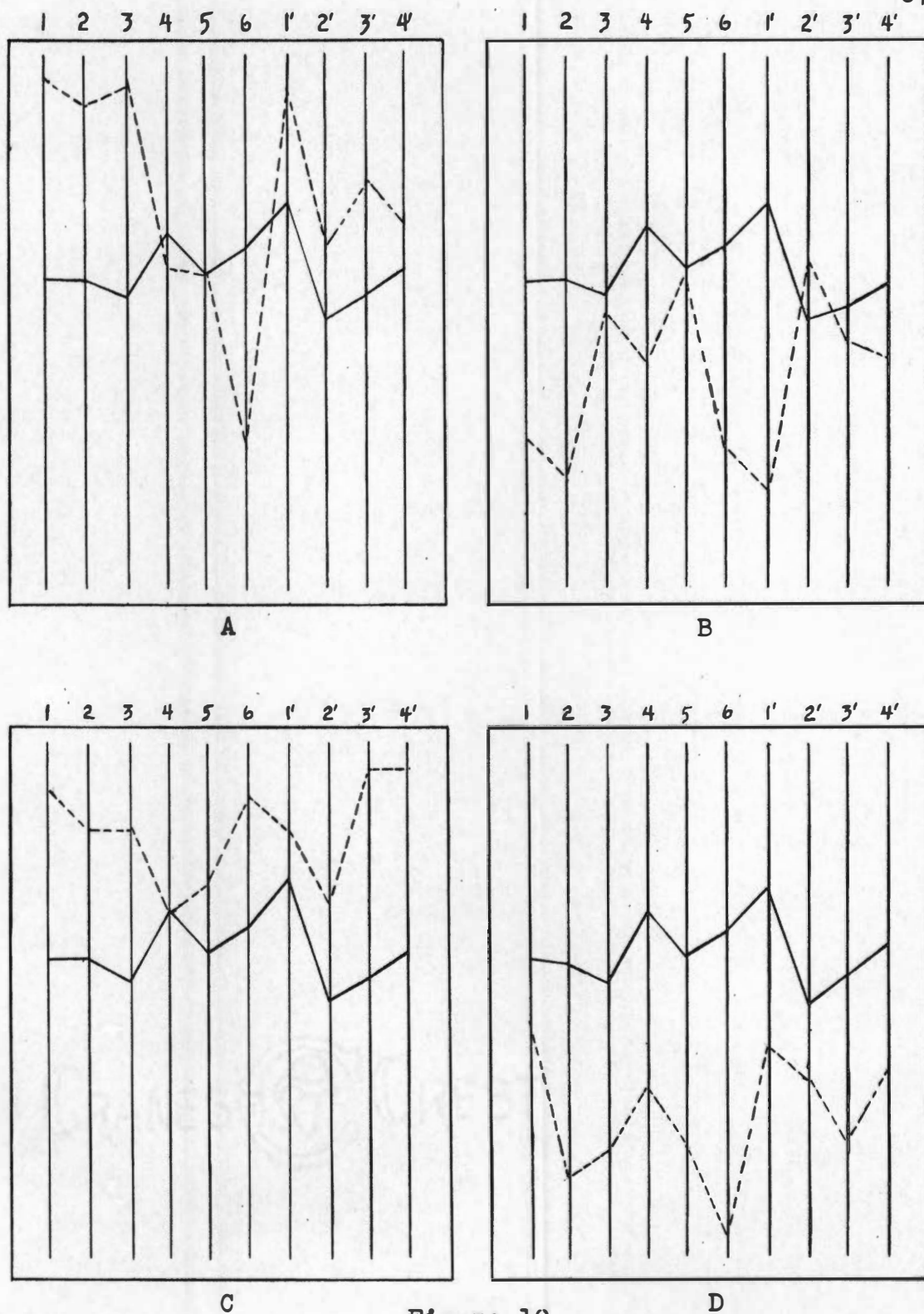


Figure 10

Samples of Individual Profiles versus Upper Group
 Profile Based on Tests Shown in Figure 9
 (Individuals A and C Represent Upper Third;
 Individuals B and D, Lower Third)

— Upper Group Mean
 - - - Individual Scores

Differences in "Total" Test Scores

A quicker method than that of plotting profile scores was devised. It required totals for a "forward" score and a "reverse" score derived from certain test scores. The tests were those used in the profile described above. No conversion of test scores was required. A description of the method follows.

The raw scores of the "forward" tests (dominance section, Edwards Personal Preference Schedule; Minnesota Teacher Attitude Inventory; Watson-Glaser Critical Thinking Appraisal; active, dominant and stable sections, Thurstone Temperament Schedule) were added in one column. The raw scores of the "reverse" tests (succorance and abasement of Edwards Personal Preference Schedule; P and M sections, P M S Scale) were added in another column. The individual who was likely to change had a high "forward" score and a low "reverse" score. Dealing with two separate scores is clumsy and may be confusing, but the two scores added arithmetically produced nothing of value according to evidence used in this study.

The "forward" score mean for the upper group was 192; the "forward" score mean for the lower group was 157. The "reverse" score mean for the upper group was 120; the "reverse" score mean for the lower group was 150. Table X shows test scores and means for the upper group. Table XI shows test scores and means for the lower group. Examination of

TABLE X

UPPER GROUP SCORES FOR CERTAIN TESTS

Upper Group												
"Forward Tests"	A	B	C	D	E	F	G	H	I	J	K	
Edwards P P S, dominance	16	9	25	18	21	25	17	10	10	24	21	
Minnesota Teacher Attitude	44	111	107	-21	64	64	93	78	75	94	13	
Watson-Glaser Critical Thinking	70	70	83	65	79	75	68	68	67	80	54	
Thurstone, active	13	13	12	14	13	12	14	14	6	13	14	
Thurstone, dominant	20	0	12	10	10	13	19	4	14	15	13	
Thurstone, stable	16	14	6	11	12	12	15	9	14	18	13	
Total Score	179	217	245	97	199	201	226	183	186	244	128	
Mean for "Forward Scores" - 192												
Mean for "Reverse Scores" - 120												
"Reverse Tests"	A	B	C	D	E	F	G	H	I	J	K	
Edwards P P S, succorance	4	20	3	5	7	10	3	15	3	5	13	
Edwards P P S, abasement	14	8	10	17	13	9	11	14	17	9	18	
P M S Schedule, P	37	31	20	38	16	20	27	39	45	11	51	
P M S Schedule, M	74	74	52	67	53	64	48	125	52	30	96	
Total Score	129	133	85	127	89	103	89	193	117	55	178	

TABLE XI

LOWER GROUP SCORES FOR CERTAIN TESTS

Lower Group											
"Forward Tests"	L	M	P	Q	R	S	T	U	V	W	Y
Edwards P P S, dominance	16	19	14	16	20	15	14	14	4	6	12
Minnesota Teacher Attitude	60	76	42	60	-20	-4	36	66	22	61	95
Watson-Glaser Critical Thinking	68	71	75	79	70	60	70	69	63	62	72
Thurstone, active	10	11	4	13	8	8	14	12	9	9	5
Thurstone, dominant	11	13	7	10	11	4	15	6	5	5	15
Thurstone, stable	15	7	19	12	13	2	16	12	6	5	16
Total Score	180	197	161	190	102	85	165	179	109	148	215
Mean for "Forward Scores" - 157											
Mean for "Reverse Scores" - 150											
"Reverse Tests"	L	M	P	Q	R	S	T	U	V	W	Y
Edwards P P S, succorance	14	22	11	13	14	14	4	18	17	16	8
Edwards P P S, abasement	19	13	18	23	9	16	9	14	21	11	15
P M S Schedule, P	38	28	50	16	43	47	58	45	51	39	49
P M S Schedule, M	74	120	63	53	79	88	82	97	56	86	69
Total Score	145	183	142	105	145	165	153	174	145	152	141

the limited data suggested that the "forward" score needs to be at least fifty greater than the "reverse" score to indicate a prediction of ability to change.

It is hypothesized by the writer that this method can be explored as one possible indicator of potential change for an individual. It is probable that further experimentation would determine whether the "forward" score or the "reverse" score would be the more valuable for prediction purposes. Also, since the individuals whose test results determined these scores were predominantly people with some teaching experience it is possible that the "total scores" could not be used for predicting the ability to change for inexperienced persons.

Tests Cited in Other Studies

Although they were not dealing with change, Luton¹² and Nunnery¹³, in separate studies, reported correlations between effective administrators (using the Rating Guide for selection)

¹²James Norfleet Luton, "A Study of the Use of Certain Standardized Tests in the Selection of Potential Educational Administrators" (Unpublished Ed. D. thesis, The University of Tennessee, March 1955), pp. 58-66.

¹³Michael Yates Nunnery, "A Study in the Use of Psychological Tests in Determining Effectiveness and Ineffectiveness Among Practicing School Administrators" (Unpublished Ed. D. thesis, The University of Tennessee, June 1958), pp. 98-105.

and certain tests or parts of tests. The present study dealt with change as indicated by the total Rating Guide and not by sections of the Rating Guide in attempting to establish relationships between certain tests, or parts of tests, and change. Both Luton and Nunnery dealt with sections of the Rating Guide as well as the total Rating Guide in attempting to differentiate between "effective" and "ineffective" groups.

According to Luton, the majority of people are "either 'effective' or 'ineffective' in most areas (of the Rating Guide) but not in all areas." He observed that "a small portion of the subjects was consistently 'effective' or consistently 'ineffective.'" The tests Luton found to be significant were summarized as follows:

The Watson-Glaser Critical Thinking test was significant in differentiating between groups judged to be unlike in the areas of behavior described under all four sections (Democratic Operation, Intelligent Operation, Adequacy of Communication, Operation as a Citizen); the Miller's Analogies and Cooperative English (Expression) were significant for three sections; the Minnesota Teacher Attitude Inventory, Cooperative English (Mechanics), "F" Scale, Kuder Preference (Music and Computational), and Allport-Vernon Study of Values (Economic and Aesthetic) were significant for two sections.¹⁴

Luton was dealing with beginning Master's students as subjects; Nunnery used practicing administrators in his study. Seeming differences in results might be partly due to lack of

¹⁴Luton, op. cit., p. 58.

refinement in differentiation between potential and practicing administrators.

The battery of psychological tests selected by Nunnery was based on tentative research findings of others. The test battery was administered to practicing school administrators; the administrators were also assigned a field rating on the Rating Guide.

It was found that (certain tests) were significant in differentiating between groups who were judged to possess different operational characteristics within the following (sections) of the Rating Guide. . . . Total Rating Guide: Allport-Vernon Study of Values (religious): middle group, low score, lower group, high score; Edwards Personal Preference Schedule (intraception): middle group, low score, lower group, high score; Edwards Personal Preference Schedule (endurance): middle group, low score, lower group, high score.

Among the battery of tests utilized in the present study there is no one test which is the best predictor of effectiveness in school administration. For those aspects of operational behavior measured by the Rating Guide the Thurstone Temperament Schedule and the Edwards Personal Preference Schedule seem to be the most promising.¹⁵

Chapter Summary

Means and percentiles (if available) of test results of the 1957 experimental group were reported in this chapter. Comparative results of the 1956 pilot experimental group were given. Some correlations between change as indicated by the

¹⁵Nunnery, op. cit., p. 105.

Rating Guide and certain test scores were suggested as bases for further experimentation. Differences in certain test scores were proposed as being significant for use as one means of predicting students' ability to change in a desirable direction. The change would be brought about through the preparation program for students in educational administration and supervision.

CHAPTER V

A STUDY OF THE 1956 PILOT GROUP TO NOTE EFFECT ONE YEAR LATER OF EXPERIMENTAL PROGRAM

Of the seventeen persons in the 1956 pilot experimental study fifteen were available for visitation or observation and rating in 1958. The individuals had been pre-rated in 1956 and post-rated in 1957 as participants in a pilot study which was reported by Taylor.¹ The third rating, made in 1958, was for the purpose of examining the development of individuals one year following their participation in the pilot program. Would the trend toward effectiveness in educational administration continue, as a group, or would it reverse?

Communication with the subjects provided for visitation dates. Rating teams who had not previously rated the individuals made the ratings. Two of the subjects had entered graduate school on a full-time basis; their ratings were based on observations in classroom and other school-related activities.

Composition of the Pilot Study Group

Among the fifteen persons, two had changed or were planning to change their graduate study emphasis from

¹Gem Kate Taylor, "The Development of Effective Characteristics in Students of School Administration" (Unpublished Ed. D. thesis, The University of Tennessee, June 1957).

educational administration to another area of education. There were six principals, three of whom were principals at the time they participated in the pilot study. Of the five who were doing regular classroom teaching, three had expressed the desire to continue classroom teaching; one had said he would like to become a coach. (He had become a part-time coach and part-time classroom teacher since 1956.) Two persons had entered the doctoral program in educational administration and supervision.

Individual Ratings

The three ratings per person are shown in Table XII. The means for each of the differences in ratings - 1956 and 1957, 1957 and 1958, 1956 and 1958 - were obtained and found to be 0.22, 0.14 and 0.36 with variances of 0.2946, 0.3197, and 0.3640 respectively. Each of these sets of paired differences was analyzed by the t-test for paired differences and only the 1956-1958 difference was found to be significant at the 0.05 level.

The four ratings for 1958 which were at or below 3.2 on the Rating Guide were made by the following persons: the two teachers whose interests had changed to another phase of education, the teacher who desired to remain in classroom teaching, and one other teacher. The ratings shown to be at or above 4.2 were made by these people: three principals, the coach-teacher and one graduate student.

TABLE XII

RATINGS OF INDIVIDUALS IN 1956 GROUP FOR THE THREE YEARS,
1956, 1957 AND 1958

Individual	Rating 1956	Rating 1957	Differ- ence	Rating 1958	Differ- ence	Differ- ence
			<u>1956</u> 1957		<u>1957</u> 1958	<u>1956</u> 1958
A	2.6	2.7	0.1	4.3	1.6	1.7
B	3.5	4.0	0.5	4.0	0.0	0.5
C	3.1	3.6	0.5	3.9	0.3	0.8
D	3.6	3.3	- 0.3	3.2	- 0.1	- 0.4
E	3.4	4.2	0.8	3.9	- 0.3	0.5
F	2.7	2.9	0.2	2.4	- 0.5	- 0.3
G	3.2	4.5	1.3	4.0	- 0.5	0.8
H	4.2	3.8	- 0.4	4.5	0.7	0.3
J	3.2	4.0	0.8	3.9	- 0.1	0.7
K	3.2	3.5	0.3	3.1	- 0.4	- 0.1
L	4.0	4.0	0.0	4.0	0.0	0.0
M	2.8	2.5	- 0.3	2.6	0.1	- 0.2
N	4.4	3.6	- 0.8	4.2	0.6	- 0.2
P	4.2	4.3	0.1	4.3	0.0	0.1
Q	3.2	3.7	0.5	4.4	0.7	1.2
Total difference			3.3		2.1	5.4
Mean difference			0.220		0.140	0.360*

*Mean difference was significant at the 0.05 level.

Group Ratings

The average 1956 rating for the fifteen people was 3.4; the average 1957 rating was 3.6; and the average 1958 rating was 3.8. The average ratings for those who were principals when they entered the program and who continued in these positions were as follows: 3.4 for 1956; 3.7 for 1957 and 4.2 for 1958. The average ratings for those who were teachers in 1956 and who since became principals were: 3.3 for 1956; 3.9 for 1957; and 4.0 for 1958.

The average ratings of those who continued to express an interest in administration but who did not hold positions in administration in 1958 were: 3.4 in 1956, 3.6 in 1957, and 3.7 in 1958. Those whose interest changed from administration to another field in education rated as follows: 3.0 in 1956, 2.9 in 1957 and 2.8 in 1958. Those who had entered graduate school between the 1956 ratings and the 1958 ratings had average ratings of 4.1 in 1956, 4.2 in 1957 and 4.2 in 1958.

Ratings by Rating Guide Sections

Ratings for three years, 1956, 1957, and 1958, are shown in Table XIII. The group showed continued improvement in three areas: Interpersonal Relations, Intelligent Operation, and Adequacy of Communication. No growth sensitive to Rating Guide measurement took place in the areas in which the

TABLE XIII

1956 GROUP AVERAGE BY RATING GUIDE SECTIONS
FOR 1956, 1957, AND 1958

Rating Guide Sections	1956	1957	1958
Interpersonal Relations, Section I	3.4	3.5	3.6
Intelligent Operation Section II	3.2	3.6	3.8
Emotional Stability Section III	4.0	4.0	4.0
Ethical and Moral Strength Section IV	4.4	3.9	4.1
Adequacy of Communication Section V	3.5	3.9	4.0
Operation as a Citizen, Section VI	3.2	3.5	3.5

Note:

Section I is composed of seven categories.

Section II is composed of eight categories.

Section III is composed of one category.

Section IV is composed of two categories.

Section V is composed of four categories.

Section VI is composed of three categories.

The Total Rating Guide is composed of twenty-five categories.

group rated high at the beginning of the study: the rating for Emotional Stability remained at 4.0, and the rating for Ethical and Moral Strength fluctuated from 4.4 to 3.9 to 4.1. The rating for Operation as a Citizen began to rise but leveled off at 3.5.

The ratings for the Total Rating Guide as given in Table XI did not appear to be an average of the ratings for the Sections; the apparent discrepancy was due to the fact that the number of items per section ranged from one to eight.

General Evaluation

Results of a control group study that was begun in the summer of 1958 may affect the accuracy of the following statement but until such time it may be said that evidence suggests that a specific kind of program in educational administration and supervision can contribute to effective change in an individual whose interest lies in the area of educational administration. This kind of program is based on the premise that a prerequisite to real change in behavior is an understanding of one's present characteristics in relation to the demands for effective job performance.

Perhaps the persons whose interest shifted from administration to other areas of education were helped to understand themselves in relation to present and future interests. Perhaps such a statement is stretching the data too far. Certainly

other circumstances may have exerted more pressures and had more influence of a lasting quality than the effect of the experimental program.

Change in the 1956 group between the first and second ratings was reported by Taylor.² The question as to whether change is permanent was only partly answered in the present study. Based on the series of three ratings, as were listed in Table IX, the three possible things which could happen did happen; that is, some individuals continued in the direction they had started, some leveled off after the first year, and some reversed direction. As a group, however, two important things took place: (1) a separation occurred between those who were interested in administration (as a group) and those who were not interested (as a group), and (2) the average trend for those interested in administration (as a group) was continuous and upward. Those not interested in administration (as a group) showed fluctuation, although also in an upward direction over the three-rating period. The Rating Guide averages for the three years of those interested in administration were: 3.5 in 1956, 3.8 in 1957 and 4.0 in 1958. The averages for each of the three years for those not interested in administration were: 3.3 in 1956, 3.1 in 1957, and 3.6 in 1958.

²Loc. cit.

Chapter Summary

A description of the 1958 composition of the 1956 pilot group was given. A table showed ratings for 1956, 1957, and 1958 of the individuals who had remained in the education profession. Changes in group rating and an analysis of the group were given. Ratings by sections and as a whole, per the Rating Guide, were given. The chapter was concluded with a summary of observations concerning the effect one year later of the program which was designed and administered as a pilot study in 1956 in the Department of Educational Administration and Supervision at the University of Tennessee.

CHAPTER VI

FINDINGS, CONCLUSIONS, AND RECOMMENDATIONS

An experimental program in educational administration and supervision was observed, analyzed and reported in terms of certain behavioral characteristics in this study. Preceding chapters have described the purpose, implementation, and outcomes of the program. This chapter presents a summary, conclusions, and recommendations for additional research.

The purpose of the experimental program was not only to teach theories and skills needed for educational administration, but to attempt to provide opportunities for the development of personal insight in relation to oneself as an educational administrator. The design of the program provided for alternating the emphasis between self-understanding and the basic skills of educational administration, plus interwoven experiences that could have borne either label.

Teaching-learning activities included psychological and communication skills, tests, speech-critique exercises, use of observers, diaries, role-playing, small and large group discussions, case studies, lectures, audio-visual aids, guest speakers and panels. Both the psychology professor and the educational administration professor participated appropriately in all phases of the program.

Summary of Findings

Behavioral Changes (as a Group) Indicated by Pre-rating and Post-rating Comparison

Comparison of pre-ratings and post-ratings of the group indicated growth in the group as a whole as measured by the total Rating Guide. By sections the amount of change varied.

Section I, Interpersonal Relations. An increased awareness of the importance of recognizing the abilities and the worth of other individuals was indicated. The group showed an increased willingness to allow participation in decision-making of those who would be affected by the decision.

Section II, Intelligent Operation. Growth in the direction of using the problem-solving approach to questions was shown. An increased willingness to make decisions on the basis of available and appropriate data rather than on the basis of a preconceived notion or tradition was indicated.

Section III, Emotional Stability. Although the group had rated high at the beginning of the program it showed growth in emotional stability. The group tended to be calm and to help others in emotion-charged situations.

Section IV, Ethical and Moral Strength. In this two-category section there were two kinds of change by the group: one change indicated growth in "having the courage of one's convictions"; the other change indicated a decreased conscientiousness about the moral responsibility for "agreements with others."

Section V. Adequacy of Communication. At the beginning of the program the group rated high in its adequacy of communication; it showed no improvement in these skills. The ratings indicated that the group was skilled in a "practical" sort of speech. Little effort seemed to be expended for increasing professional leadership in the area of communication.

Section VI, Operation as a Citizen. The pre-rating and post-rating were similar and indicated little more than average community leadership. There was little evidence of active interest in the affairs of citizenship, although some "intellectual" interest was indicated.

Comparison of Change in This Group with Change in an Earlier Group

Changes which occurred during the eight-to-ten months period between ratings in the experimental group of 1957 and in the pilot experimental group of 1956 were compared. The pilot group had been composed entirely of men; the experimental group of 1957 was one-third women. The groups were similar in respect to age, experience and positions. Seventeen persons were involved in the first study; eighteen participated throughout the second study.

Both groups showed growth as indicated by change in total Rating Guide ratings. Both groups seemed to change with respect to characteristics involving Interpersonal Relations, Intelligent Operation, and Emotional Stability.

The 1956 group showed change in the areas of Operation as a Citizen and Adequacy of Communication. In partial contrast to the 1957 group, the 1956 group showed behavioral growth in all areas of the Rating Guide except one, whereas the 1957 group showed growth in three of the six areas.

Description of the 1957 Group as Indicated by Tests

Tests were administered to the group early in the program. The tests were chiefly for the purpose of increasing self-understanding; the other purpose was to attempt to identify differing characteristics between individuals who changed and those who did not change. Results of the tests on a total group basis follow.

Allport-Vernon Scale of Values. The test was designed to measure the relative prominence of six basic motives in personality, theoretical, economic, aesthetic, social, political and religious. The group fell within the "average" range for all values except just below the lower limit of "average" for aesthetic.

Edwards Personal Preference Schedule. The P P S was designed to measure a number of relatively independent normal personality variables. The only mean score outside the "average" range was a "low" rating on heterosexuality by men. This suggests that their needs in relation to members of the opposite sex were being met more than adequately; or perhaps that they were too self-conscious or shy to express their true

feelings when they answered questions on the Schedule.

Thurstone Temperament Schedule. The test was devised for the purpose of measuring important stable traits which describe how normal people differ from each other. The group fell within the "average" range with the exception of the women in two areas: they rated "low" on impulsive and "high" on stable. A low score on impulsive indicates a person who stands by his convictions and refuses to be swayed by unreasonable argument. A high stable score indicates an even disposition and a calmness in critical situations.

Cooperative English. The mechanics section measured the degree of mastery of the essentials of correct English usage. The effectiveness of expression section measured the skills of sentence structure and style, active vocabulary, and organization. The means of the group for both sections were similar to means for seniors in colleges that were "representative of many junior and teachers colleges."

Minnesota Teacher Attitude Inventory. The Minnesota Teacher Attitude Inventory was administered at the beginning and at the end of the summer program. This test lent itself to verbal change. The group scored a little lower than the pilot group on the first test and slightly higher than the pilot group on the second test.

P M S Scale. The scale measured the degree to which one employs magical thinking. Three phases relating to

magical thinking were measured. The group scored below the standardized mean in the areas of "power" and "magic"; the group mean was the same as the standardized mean for "superstition."

Professional Beliefs Inventory. Two portions of this inventory were administered to the group. The first dealt with Child Growth and Development; the other with The School in a Democratic Society. No normative means were available. The scores were similar to those of the 1956 group.

Correlation of Change (as Indicated by the Rating Guide) and Test Results

Ten tests or parts of tests suggested that further experimentation might establish the existence of significant correlations between test scores and change, as indicated by the Rating Guide. Tests or sections of tests which tended to show a positive correlation were: dominance section, Edwards Personal Preference Schedule; Minnesota Teacher Attitude Inventory; Watson-Glaser Critical Thinking Appraisal; active, dominant and stable sections, Thurstone Temperament Schedule. Sections of tests that tended to show a negative correlation were: succorance and abasement sections, Edwards Personal Preference Schedule; and the P and M sections P M S Scale.

Uses of the Tests for Predicting Change

Means for certain tests were distinctly different for the "upper third" and the "lower third." The "upper third"

was defined as "those whose post-ratings fell above the group post-rating mean and who fell in the upper third in regard to amount of total change as indicated by pre-rating and post-rating comparison." The "lower third" was defined as "those whose post-ratings fell below the group post-rating mean and who fell in the bottom third in regard to amount of total change, as measured by differences in pre-ratings and post-ratings." Chapter IV suggested two ways the test mean differences might be used for predicting change on a group basis for students who have had teaching experience.

Observation of Post-Experiment Effect of Program

Individuals showed varied patterns of change over the period of three observations. By the time of the third observation the group had become sharply defined into two divisions; namely, those who definitely wished to become administrators or supervisors and those who wished to continue in present non-administrative positions or to enter areas of education other than administration or supervision. The first group showed continued upward growth; the second group fluctuated but was moving upward at the third rating.

Conclusions

An attempt was made to assay findings of a pilot study of 1956 and to analyze certain data pertaining to the participants in the 1956 and the 1957 experimental programs for clues

to personality characteristics which affect change. The hypothesis for the study was that methods could be employed for discerning and measuring changes in certain characteristics of beginning Master's students, and that distinguishing characteristics between those who change and those who do not change could be, at least grossly, defined. The hypothesis was substantiated to some degree as will be indicated below.

The pilot study had endeavored to determine the characteristics for effective administration which "seem to change or improve in a program designed to aid the student in achieving an understanding of his personal characteristics" in relation to his job performance.¹ That study emphasized the importance of conducting such a program with emphasis on "a learning environment that provides an atmosphere which stimulates responsibility for learners to seek appropriate values, and a process for the solution of problems."² The concerns expressed for the pilot program were of vital importance in the second program. Some methods were altered; some results were different. The experimentation involved individuals, with the variations that constitute individuality.

¹Gem Kate Taylor, "The Development of Effective Characteristics in Students of School Administration" (Unpublished Ed. D. thesis, The University of Tennessee, June 1957), p. 111.

²Loc. cit.

Comparison with Findings of the Pilot Study

Characteristics in three areas of behavior described in the Rating Guide did change in both groups. The areas were: Interpersonal Relations, Intelligent Operation, and Emotional Stability. Such change seemed to indicate the effectiveness of the portions of the program designed to produce change in those areas.

The pilot study had reported changes in characteristics in the areas of Adequacy of Communication and Operation as a Citizen. Two possible reasons for the lack of improvement in the area of communication for the second group may have been: (1) the group had a high rating at the beginning; (2) less emphasis was put on communication skills in the second program. The only reason that presented itself for the lack of development in the area of citizenship was that the program depended on the incidental development of characteristics in that area; that is, the role of a school administrator as a citizen was not sharply defined.

Neither group showed growth in the total area of Ethical and Moral Strength; however, both showed growth in the category that dealt with "having the courage of one's convictions." Also, both groups rated high in this area.

Characteristics Indicative of Ability to Change

Certain tests results were different, on a group basis, for individuals who changed between pre-ratings and post-ratings

and those who did not change. As a group those who changed had high scores on the following tests or sections of tests: dominance, Edwards Personal Preference Schedule; Minnesota Teacher Attitude Inventory; Watson-Glaser Critical Thinking Test; active, dominant, and stable, Thurstone Temperament Schedule. On the same tests those who did not change scored low, as a group.

Those who changed had low scores (as a group) on the following test sections: succorance and abasement, Edwards Personal Preference Schedule; and P and M, P M S Schedule. Those who did not change scored high (as a group) on those tests. The number of persons tested in the combined experimental groups was small (test scores of thirty-three individuals were used in the analyses reported in this investigation), but the findings may become a source for further research.

Recommendations

Validation of prediction methods. It is recommended that data concerning students who were involved in two experiments in the Department of Educational Administration and Supervision at the University of Tennessee in the summer of 1958 be subjected to examination to help validate or invalidate the correlations suggested in this study. Such correlation would be helpful in the selection of students on a group

basis. Perhaps degree of change could become part of the prediction.

Development of techniques for effecting change. If the area of Operation as a Citizen remains an important section of the Tennessee Rating Guide, it is recommended that research be undertaken for the purpose of developing or defining techniques that produce change in individuals in this area. Although the evidence was inconclusive about change in the area of Ethical and Moral Strength, techniques for bringing about change in this area might be improved. (Both groups observed were behaving at a high level in this area, however, and unmeasured change may have taken place.)

Refinement of upper limits of Rating Guide. Perhaps the Rating Guide could be divided into more sensitive graduations in the upper levels. The Guide seemed to be highly effective in measuring differences between the "effective" and the "ineffective" potential and practicing administrators. A separate study now under way at the University of Tennessee may determine well-defined differences for the "average" administrator. But the measurement of growth in the 1956 and 1957 experimental groups appeared to level off close to the four rating. Refinement between the four and five ratings would probably help indicate whether or not change was actually occurring at the high levels.

Incorporation of experimental program design into regular academic program. The writer recommends that a modified

program similar to the experimental programs be designed as a part of the regular academic program. Details might require a series of courses with a revised distribution of credits. The verbal praise of most of the participants in the 1956 and 1957 experiments amounted to an enthusiastic endorsement of the experimental programs.



BIBLIOGRAPHY

BIBLIOGRAPHY

- Allport, Gordon W., Phillip E. Vernon and Gardner Lindzey. Study of Values. Boston: Houghton Mifflin Company, 1951. (Manual of Direction)
- Cook, Walter W. Minnesota Teacher Attitude Inventory. New York: The Psychological Corporation, 1951. (Manual of Directions)
- Cooperative English Test. Percentile Ranks for High School and College Students. New York: The Cooperative Test Service, 1940. (Manual)
- Edwards, Allen L. Edwards Personal Preference Schedule. New York: The Psychological Corporation, 1954. (Manual of Directions)
- Evernden, William Lyle. "The Reliability of the Tennessee Rating Guide for School Administrators." Unpublished Master's thesis, The University of Tennessee, June 1958.
- Gentry, Harold Wayne. "Patterns of Behavioral Characteristics Exhibited by School Administrators." Unpublished Ed. D. thesis, The University of Tennessee, March 1957.
- Graff, Orin B. and Ralph B. Kimbrough. "What We Have Learned about Selection," Phi Delta Kappan, 38 (April 1956), 295.
- Greever, Clarence E. "A Study of the Characteristics of Selected Effective Superintendencies in East Tennessee." Unpublished Ed. D. thesis, The University of Tennessee, June 1956.
- Jahoda, Marie, Morton Deutsch and Stuart W. Cook. Research Methods in Public Relations, Part One: Basic Processes. New York: The Dryden Press, 1951.
- _____, Research Methods in Public Relations, Part Two: Selected Techniques. New York: The Dryden Press, 1951.
- Luton, James N. "A Study of the Use of Certain Standardized Tests in the Selection of Potential Educational Administrators." Unpublished Ed. D. thesis, The University of Tennessee, March 1955.
- Moffett, Charles R. "Operational Characteristics of Beginning Master's Students in Educational Administration and Supervision." Unpublished Ed. D. thesis, The University of Tennessee, June 1954.

- Nunnery, Michael Yates. "A Study in the Use of Psychological Tests in Determining Effectiveness and Ineffectiveness among Practicing School Administrators." Unpublished Ed. D. thesis, The University of Tennessee, June 1958.
- Schmitt, Leonard R. "The Behavioral Characteristics of School Principals." Unpublished Ed. D. thesis, The University of Tennessee, June 1958.
- Southern States Cooperative Program in Educational Administration. Better Teaching in School Administration. Nashville: SSCPEA, 1955.
- Taba, Hilda et al. Diagnosing Human Relations Needs. Washington: American Council on Education, 1955.
- Taylor, Gem Kate. "The Development of Effective Characteristics in Students of School Administration." Unpublished Ed. D. thesis, The University of Tennessee, June 1957.
- Thurstone, L. L. Thurstone Temperament Schedule. Chicago: Science Research Associates, Inc., 1953. (Manual of Directions)
- Watson, Goodwin and Edward Maynard Glaser. Watson-Glaser Critical Thinking Appraisal. Yonkers-on-Hudson: World Book Company, 1952. (Manual of Directions)

APPENDIX

CRATES OF CREST

UNIVERSITY OF TENNESSEE RATING GUIDE

Characteristics of School Administrators

As Revised January 8, 1958

I. Interpersonal Relations

A. How does he relate to others?

- ☐ 1. Tends to be a lone wolf
- ☐ 2. Has a few friends but tends to ignore others
- ☐ 3. Friendly when approached by others
- ☐ 4. Popular; has many casual acquaintances
- ☐ 5. Steadily warm and appealing in relationship with others

B. Does he utilize the opinion of others?

- ☐ 1. Generally ignores the viewpoints of others
- ☐ 2. Uses opinions if they agree with his own
- ☐ 3. Values opinions of those who volunteer suggestions but fails to seek opinions of others
- ☐ 4. Highly selective in utilizing opinions; sometimes values ideas that differ from his own
- ☐ 5. Consistently seeks and considers the opinions of others

C. Is he skillful in developing an organization in which each can do his best?

- ☐ 1. Most people with whom he works have important responsibilities in which they are genuinely interested
- ☐ 2. Sometimes delegates responsibilities with regard to special interests and abilities of associates
- ☐ 3. Delegates tasks largely mechanically; fails to recognize special abilities of others
- ☐ 4. Plays favorites in delegating responsibility
- ☐ 5. Runs the whole show himself

D. Is he skillful in getting policies formulated co-operatively?

- ☐ 1. Involves general public, staff members and students in major policy formulation
- ☐ 2. Attempts to involve general public, staff members and students in policy formulation but has difficulty in setting up necessary machinery
- ☐ 3. Involves only key people in policy formulation
- ☐ 4. Discusses policies with others, but decisions are usually made prior to the discussion
- ☐ 5. Formulates policies himself; rarely discusses them with others

E. Is he skillful in continuous implementation of policies?

- ☐ 1. Tends to ignore or defer action on policies
- ☐ 2. Vacillates in implementing policies
- ☐ 3. Tends to force policies without sufficient planning
- ☐ 4. Cautious in effecting policies
- ☐ 5. Moves surely and judiciously in effecting policies

F. Does he help the group arrive at a working consensus?

- ☐ 1. Contributes little to help group arrive at a working consensus
- ☐ 2. Tries to force group to quick agreements without really considering problems
- ☐ 3. Tends to force action without careful group consideration
- ☐ 4. Strives for consensus but sometimes encourages group action on insufficient data
- ☐ 5. Continually strives for careful group problem analysis; helps group recognize points of agreement

G. Does he believe that democratic processes are essential?

- ☐ 1. Urges the use of processes consistent with best democratic practices
- ☐ 2. Is cognizant of responsibility to use democratic procedures; is sometimes unsure of how to employ them
- ☐ 3. Attempts to use democratic processes; however, usually resorts to expediency in pressing situations
- ☐ 4. Gives lip service to democratic processes which are not evident in his behavior
- ☐ 5. Uses any expedient method available to attain a pre-determined end

II. Intelligent Operation

A. Does he give sufficient consideration to new data in problem solving?

- ☐ 1. Disregards new data that challenge the status quo
- ☐ 2. Uses new data only when they support his position
- ☐ 3. Will consider new data when presented to him
- ☐ 4. Seeks new data along lines of special interests
- ☐ 5. Consistently seeks and employs new data

B. Does he recognize and define problems?

- ☐ 1. Tends not to recognize the existence of problems
- ☐ 2. Tends to consider symptoms instead of problems
- ☐ 3. Sometimes confuses symptoms with problems in his efforts to improve
- ☐ 4. Recognizes problems but has difficulty in analyzing them
- ☐ 5. Recognizes and analyzes problems

C. Is he consistent in terms of his basic assumptions?

- ☐ 1. Supports conflicting ideas; action characterized by inconsistency
- ☐ 2. Has a tendency to discuss important problems in terms of his likes and dislikes
- ☐ 3. Frequently uncertain of his position on controversial subjects
- ☐ 4. Is certain of his position and consistent in his behavior in areas which he considers important
- ☐ 5. Is dependable and predictable in word and action

D. Does he experiment?

- ☐ 1. Tends to try out new ideas after careful study and follows through on basis of experimental evidence
- ☐ 2. Undertakes various new projects for improvement but fails to interpret their significance
- ☐ 3. May be premature in trying out ideas for improvement; fails to fully incorporate accepted principles of experimentation
- ☐ 4. Action tends to be based on subjective evidence
- ☐ 5. Tends to operate within traditional practices or on basis of hunches

E. Does he try to recognize and deal with his own biases?

- ___ 1. Consistently examines his own position and attempts to understand the position of others
- ___ 2. Tends to evaluate his position but will resort to biases under pressure
- ___ 3. Feels uneasy about his position at times; can be stimulated to examine his opinions
- ___ 4. Assumes that his position is generally right; does not know how to identify his own biases
- ___ 5. Refuses to examine his position

F. Does he appear to have profited by previous experience?

- ___ 1. Frequently makes the same mistake but seldom admits it
- ___ 2. Usually attempts to justify mistakes
- ___ 3. Recognizes that some mistakes are inevitable but has difficulty in making readjustments
- ___ 4. Makes some improvement as a result of past mistakes
- ___ 5. Recognizes his mistakes and seeks to avoid repeating them

G. Does he have the ability to size up people

- ___ 1. Judges potentialities of people in terms of their race, religion, nationality, or other such concepts
- ___ 2. Makes judgments about people in terms of hunches
- ___ 3. Tends to base judgments of people on past experiences without rethinking in terms of present situations
- ___ 4. Judges people on basis of personal experiences, using additional resources when problem situations arise
- ___ 5. Consciously endeavors to understand the basic potentialities of each person through objective procedures

H. Does he accept responsibility wisely?

- ☐ 1. Budgets the assuming of responsibilities wisely in terms of own limitations in present situation
- ☐ 2. Carries out pressing responsibilities well but neglects less urgent duties
- ☐ 3. Concentrates on school routine; supports non-school endeavors on a highly selective basis
- ☐ 4. Attends strictly to school routine without participating in community enterprises
- ☐ 5. Accepts too many responsibilities or refuses to assume responsibilities normally expected of him

III. Emotional Stability

A. Is he emotionally stable?

- ☐ 1. Tends to be upset by everyday occurrences and keeps staff in continuous uproar
- ☐ 2. Attempts to exemplify outward calmness but explodes about trivial matters
- ☐ 3. Is upset in novel situations and has a tendency to upset others
- ☐ 4. Meets novel situations well but lets some problems involve him in distracting entanglements
- ☐ 5. Appears to meet crises with a contagious calmness; others feel at ease in his presence

IV. Ethical and Moral Strength

A. Does he have the courage of his convictions?

- ☐ 1. Tends to weasel out of situations
- ☐ 2. Usually follows most popular viewpoint
- ☐ 3. Has a tendency to accept some viewpoints which he realizes are in conflict with his own
- ☐ 4. Has well-tempered convictions which he tries to follow but is sometimes unsure of their soundness
- ☐ 5. Places principle above his own personal advantage

B. Does he exhibit integrity in dealing with others?

- ☐ 1. Considers agreements with others as promissory notes to which he is committed
- ☐ 2. Exhibits integrity in important agreements, but in less important agreements is somewhat careless
- ☐ 3. Tends to rationalize inadvertent breaches of agreements
- ☐ 4. Through indirect methods leads people to believe in false situations
- ☐ 5. Tends to be unscrupulous in accomplishing his purposes

V. Adequacy of Communication

A. How well does he express himself orally?

- ☐ 1. Chooses words which clearly convey thoughts; is able to express abstract ideas
- ☐ 2. Expresses practical thoughts fairly well, but has difficulty with abstractions
- ☐ 3. Is unimpressive in oral communication
- ☐ 4. Expresses himself in a fuzzy, incomprehensible manner and tends to puzzle listeners concerning what he means
- ☐ 5. Is either unable or does not desire to convey thoughts to others

B. Is he a good listener?

- ☐ 1. Is attentive in trying to grasp ideas expressed by others
- ☐ 2. Listens carefully to ideas in which he is interested
- ☐ 3. Appears to listen but has difficulty in concentration
- ☐ 4. Tends to disrupt oral communication by inattentiveness or by introduction of irrelevant ideas
- ☐ 5. Tends to listen only to himself

C. Does he interest people in examining ideas?

- ☐ 1. Stimulates people to seek solutions through critical analyses of ideas
- ☐ 2. Encourages examination of ideas that he thinks are important
- ☐ 3. Waxes hot and cold in stimulating examination of ideas
- ☐ 4. Appears to consider intellectual curiosity unimportant
- ☐ 5. Discourages examination of ideas

D. How skillfully does he lead group discussions?

- ☐ 1. Is either at a loss or monopolizes discussion when appointed official leader of a group
- ☐ 2. Permits everyone to talk without achieving a group decision
- ☐ 3. Tends to rely on key persons in group discussions
- ☐ 4. Operates well within a structured agenda
- ☐ 5. Facilitates a stimulating and well-ordered climate conducive to reaching group decisions

VI. Operation as a Citizen

A. Does he help people interpret significant contemporary trends and events?

- ☐ 1. Does not seem to be informed about or interested in contemporary events
- ☐ 2. Discusses current affairs in terms of stock phases and generalities
- ☐ 3. Knows about current affairs but shows prejudice in discussing them
- ☐ 4. Is well informed in the socioeconomic problems in which he is interested
- ☐ 5. Discusses intelligently major social, political and economic issues

B. Is he cooperative with non-educational groups working for community betterment?

- ___ 1. Is aware of and actively concerned with desires and interests of community groups, agencies and organizations
- ___ 2. Is interested in cooperating with community groups but spreads his efforts too thinly
- ___ 3. Is selective in cooperating with groups in proportion to pressures applied
- ___ 4. Becomes so involved with activities of non-educational groups that he neglects proper administration of the school program
- ___ 5. Considers the school an island that is competitive with non-educational groups

C. What is his attitude toward minority groups in the school community?

- ___ 1. Insists that minority points of view be appropriately represented in community-school decisions
- ___ 2. Upholds right of most minority viewpoints to be represented but neglects consideration of those that are extreme
- ___ 3. Follows a hands-off policy in regard to minority groups in the community
- ___ 4. Tends to ignore the existence of minority groups in the community
- ___ 5. Indicates that minority groups have no right to representation in community-school affairs