A STUDY OF COMPUTER SCORED GROUP HOLTZMAN INKBLOT VARIABLES AS RELATED TO STUDENT TEACHING SUCCESS, MAJOR TEACHING FIELDS, AND SEX

Thesis for the Degree of Ph. D.

MICHIGAN STATE UNIVERSITY

B. BRADLEY WEST

1969

This is to certify that the

thesis entitled

A STUDY OF COMPUTER SCORED GROUP HOLTZMAN INKBLOT VARIABLES AS RELATED TO STUDENT TEACHING SUCCESS, MAJOR TEACHING FIELDS, AND SEX

presented by

B. Bradley West

has been accepted towards fulfillment of the requirements for

Ph.D. **degree in** Education

W. Henry Kennedy

Major professor

Date November 4, 1969

ABSTRACT

A STUDY OF COMPUTER SCORED GROUP HOLTZMAN INKBLOT VARIABLES AS RELATED TO STUDENT TEACHING SUCCESS, MAJOR TEACHING FIELDS, AND SEX

By

B. Bradley West

Purpose of the Study

For years researchers have attempted to identify the personality correlates of successful teachers. This study sought to deal with this problem in terms of perceptual measures. Specifically, the purpose of the study was to determine the relationship between measures of perception as identified by Holtzman Inkblot variable scores and success in student teaching. Sub-problems studied the relationship between perceptual measures and student teachers grouped according to sex and major teaching field.

Methodology

The group version of the Holtzman Inkblot Test was administered to 802 Michigan State University student teachers during Winter Term, 1968, in each of the geographical centers where students are assigned for student teaching. Forty-five inkblot slides were exposed to all the student teachers for one minute each and the subjects

responded with a short written description of what each inkblot "looked like" on the specially designed group answer sheet. The answer sheets and identifying data were keypunched, loaded onto magnetic tape and computer scored by the Holtzman Inkblot Scoring Service in Perry Point, Maryland. Output from the Scoring Service included a print out and a score summary card for each subject in the sample on seventeen Holtzman variables: Location, Rejection, Form Definiteness, Color, Shading, Movement, Integration, Human, Animal, Anatomy, Sex, Abstract, Anxiety, Hostility, Barrier, Penetration and Popular.

Findings of the Study

When the sample of student teachers was grouped on the basis of success ratings at the conclusion of their student teaching, a study of the mean Holtzman scores for the above seventeen variables showed that only one variable score, Rejection, was significantly different among the groups. But the unequal variance of the R variable among the groups suggested that even this single significant finding was questionable. A re-grouping of the student teachers by sex and then success ratings failed to identify any other significantly different variable scores. When the student teachers were grouped by sex alone with no additional sub-grouping, eleven of the

seventeen variable scores significantly differentiated between males and females. Thus, the Holtzman variable scores would appear to be of some usefulness in discriminating between male and female student teachers and suggested the importance of analyzing the data for the sex influence. When the sample of student teachers was grouped into thirteen major teaching areas, nine of the seventeen variables differentiated among the various teaching areas. However, when the sample was first grouped according to sex and then sub-grouped into major teaching area, one variable significantly differentiated among the males while eleven variables were significantly different among females. Thus, the Holtzman scores would appear to be of little or no value in discriminating among male student teachers grouped according to major teaching fields but would be of some, but restricted, value when applied to females grouped according to major teaching area.

The study also generated perceptual norms on the Holtzman for Michigan State University student teachers and determined the feasibility of administering and computer scoring a projective technique with a large (N = 802) population.

A STUDY OF COMPUTER SCORED GROUP HOLTZMAN INKBLOT VARIABLES AS RELATED TO STUDENT TEACHING SUCCESS, MAJOR TEACHING FIELDS, AND SEX

By

B. Bradley West

A THESIS

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

Student Teaching Office

1969

ACKNOWLEDGMENTS

The investigator wishes to express his gratitude and thanks to the many colleagues, faculty members, students and friends who gave their time, cooperation, and effort to this study.

The experimenter is indebted to Dr. W. Henry Kennedy, Committee Chairman and Dissertation Director, for his counsel, guidance, thought provoking questions and encouragement as the study was conceived and undertaken. With his scholarly guidance and practical advice, the study was strengthened and considerably improved.

To Dr. Donald Gorham, Chief, Research Psychology, at the Veterans Administration Hospital in Perry Point, Maryland, the author offers special thanks for his interest, help and cooperation in computer scoring the hundreds of test protocols;

To my wife, Dorothy, for her support, critical reading and unfailing encouragement to get the job done;

To Dr. and Mrs. Kent Gustafson who contributed significantly at the initial planning stages;

To Dr. Albert Rabin, who first suggested the computer-scoring approach;

To Dr. Wayne H. Holtzman, Dean, College of Education, University of Texas, for providing the inkblot slides;

To Drs. George Myers, John Suehr, and Arthur Seagull, members of the Guidance Committee, for their interest and critical evaluation of the dissertation;

To the Student Teaching Coordinators for their splendid cooperation in arranging testing dates, times and locations;

To The College of Educational Research Bureau and particularly to Mr. David Wright for assistance in the statistical analysis;

To the staff of the Computer Laboratory who spent many hours loading the 40,000 keypunched cards onto magnetic tape;

To The Student Teaching Office Graduate Affairs Committee for financial assistance in conducting the research;

And finally, to the hundreds of unnamed student teachers who willingly took part in the study;

To all of these, the investigator expresses his sincerest appreciation and gratitude.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF APPENDICES

CHAPTER I

INTRODUCTION

Statement of the Problem

The major problem of this study was to determine the relationship between computer scored group Holtzman Inkblot variables and success in student teaching. Sub-problems included study of the relationships between Holtzman variables, and major teaching fields, and sex of the student teachers.

Significance of the Study

For more than a generation educational researchers have attempted to identify the personality correlates of successful teaching behavior. The question of how to describe the successful teacher personality has been the object of hundreds of research inquiries. Contemporary literature in personality and teaching success provided ample support for the belief that personality factors contribute significantly to success in teaching. Getzels and Jackson, for example, reported conclusions in a major compilation of teacher personality research studies, and stated that, "the personality of the teacher is a significant variable in the classroom. Indeed, some would

argue that it is the most significant variable."[1] Symonds, a prolific writer and researcher on teacher effectiveness, reported, "it is now well known that personality is an important factor in the effectiveness of the teacher."[2] Arthur Combs, a leading spokesman in teacher education, summed up the literature when he wrote, "A good teacher is primarily a personality."[3]

Administrators of teacher education programs also accept personality as an important factor in admission to programs. Studies by Stout[4] and Magee[5] confirmed that teacher training program administrators consider personality the single most important selection factor; Willcox and Beigel[6] wrote that such measures could serve

---

[1]J. W. Getzels and P. W. Jackson, "The Teacher's Personality and Characteristics," Handbook of Research on Teaching, ed. by N. L. Gage (Chicago: Rand McNally and Company, 1963), p. 506.

[2]Percival M. Symonds and Stephanie Dudek, "Use of the Rorschach in the Diagnosis of Teacher Effectiveness," Journal of Projective Techniques, XX (June, 1956), 227.

[3]Arthur W. Combs, "The Personal Approach to Good Teaching," Educational Leadership, XXI (March, 1964), 327.

[4]Ruth Stout, "Selective Admissions and Retention Practices in Teacher Education," Journal of Teacher Education, VIII (December, 1957), 422-432.

[5]Robert M. Magee, "Admission-Retention in Teacher Education," Journal of Teacher Education, XII (March, 1961), 81-85.

[6]Isabel Willcox and Hugo G. Beigel, "Motivations in the Choice of Teaching," Journal of Teacher Education, IV (June, 1953), 106-109.

as selection criteria for admitting candidates into teacher education programs. And the profession at large, speaking through the Teacher Education and Professional Standards Commission (TEPS) of the National Education Association recommended that,

> Screening should include observation of personality traits and emotional stability . . . and colleges should devise and use a measure of personality.[7]

Furthermore, teacher employers are reluctant to hire candidates without some specific comments about the personality factor, as Charles[8] suggested.

But there is a gap between opinion and practice. Inlow,[9] studying thirty-eight midwestern colleges offering teacher education, reported that only eleven of the thirty-eight institutions used any standard procedure of personality evaluation and, of these, all but two relied entirely upon teacher-student interviews for accomplishing their purpose.

---

[7]National Commission on Teacher Education and Professional Standards, The Education of Teachers: Considerations in Planning Institutional Programs (Washington, D.C.: National Education Association, 1960), p. 6.

[8]Harvey Charles, "The Use of a Projective Technique in Teacher Selection," National Catholic Educational Association Bulletin, LVII (August, 1961), 172-173.

[9]Gail M. Inlow, "Comparative Study of Student Teaching Practices in 38 Midwestern Institutions," Journal of Experimental Education, XXVIII (June, 1960), 337-349.

> It appears that although there is considerable opinion favoring the use of personality as a method of selection, administrators are reluctant to make further use of this approach until the body of factual data concerning teacher characteristics has been enlarged and substantiated.[10]

Barr,[11] Mitzell,[12] Getzels and Jackson,[13] and Ryans,[14] extensively reviewed recent efforts to enlarge and substantiate teacher personality data and concluded that no consensus exists on what combination of personality factors are characteristic of a successful teacher. In the studies reported by Morsh and Wilder[15] researchers attempted to investigate successful teaching using administrator ratings, peer ratings, self-ratings

---

[10]Thomas K. Allen, "Personality as a Predictor of Student Teaching Success," Journal of the Student Personnel Association for Teacher Education, VI (Fall, 1966), 12.

[11]Arvil S. Barr, "Wisconsin Studies of the Measurement and Prediction of Teacher Effectiveness: A Summary of Investigations," Journal of Experimental Education, XXX (September, 1961), 5-156.

[12]Harold E. Mitzel, "Criteria of Teacher Effectiveness," Encyclopedia of Educational Research, ed. by C. W. Harris (MacMillian Company, 1960), pp. 1481-1486.

[13]Getzels and Jackson, op. cit., pp. 506-582.

[14]David G. Ryans, "Assessment of Teacher Behavior and Instruction," Review of Educational Research, XXXIII (October, 1963), 415-1441.

[15]Joseph E. Morsh and Eleanor W. Wilder, Identifying the Effective Instructor: A Review of the Quantitative Studies, 1900-1952 (Research Bulletin, AFPTRC-TR 54:44, San Antonio: United States Air Force Personnel Training Research Center, 1954), pp. 1-151.

and student gains criteria. The researchers also attempted to relate such variables as intelligence, age, scholarship, education and a host of other variables to successful teaching.[16] But most findings were inconclusive. The studies were hampered and compounded by situation, criterion and predictor limitations, teacher shortages, theoretical issues and difficulty in constructing instruments which predict teaching success.

Thus, despite the voluminous research concerning this problem, very little evidence has emerged which might be generally useful in the selection of teacher education candidates and in future on-the-job predictions. In effect, this constituted a demand for information clarifying the relationship between personality and teacher success. Further research is needed that will lead to the discovery of specific and distinctive features of the successful teacher personality. Perhaps no <u>one</u> personality factor will ever be found to predict success in teaching, but rather a <u>pattern</u> of personality factors. As Lamke pointed out, success in teaching may require a balance of personality factors.

> Teachers are successful . . . because there is a kind of "balance" among their personality

---

[16]Simeon J. Domas and David V. Tideman, "Teacher Competence: An Annotated Bibliography," <u>Journal of Experimental Education</u>, XIX (December, 1950), 101-218.

> traits. Another kind of balance means they will probably be poor teachers. Lacking either, they will likely be average.[17]

Lamke further suggested that success prediction demands an understanding of what is required for this balance. Study of the association of traits, one by one, with success will not suffice.

Theoretical Basis of the Study

However, within the course of this prodigious effort there has appeared a small but important group of studies which utilized projective techniques to study teacher personality which suggested an interesting phenomenon. These studies[18,19,20] suggested, each in its own way, that teachers characteristically view their world--give it structure and meaning--in ways that differentiate them from non-teaching groups. Further, this

---

[17]Thomas A. Lamke, "Personality and Teaching Success," Journal of Experimental Education, XX (December, 1951), 247.

[18]Florence L. Goodenough, Elizabeth M. Fuller, and Edna Olson, "The Use of the Goodenough Speed-of-Association Test in the Pre-service Selection of Nursery School-Kindergarten-Primary Teachers," Journal of Educational Psychology, XXXVII (September, 1946), 335-346.

[19]Thomas L. Hilton, "Ego Involvement in Teaching: Its Theory and Measurement by a Word Completion Technique" (unpublished Ph.D. dissertation, Harvard University, 1955).

[20]Merle M. Ohlson and Raymond E. Schultz, "Projective Test Response Patterns for Best and Poorest Student Teachers," Education and Psychological Measurement, XV (Spring, 1955), 18-27.

perceptual phenomenon also appeared to distinguish teachers rated successful from those rated unsuccessful. The fact that students in training showed this perceptual tendency suggested, too, that it is not merely a product of experience.

Since perception is a phenomenon of personality, it was reasonable to hypothesize that the personality structure of successful teachers was significantly different from that of unsuccessful teachers. Arthur Combs wrote that,

> Whether an individual will be an effective teacher depends upon the nature of his private world of perceptions.[21]

Since direct research on personality, competencies, and other major variables has been unable to identify any common trait or practice of successful teaching, and in view of the positive indications of studying teacher personality through perception, the present investigation attempted to identify perceptual modes of student teachers classified according to (1) success rating, (2) sex and (3) major teaching field. Other possible classifications such as age, intelligence, and occupational level also seemed appropriate, but since these are common background characteristics of the sample (i.e., all in the sample were in the same approximate age group, all met academic

---

[21]Arthur W. Combs, The Professional Education of Teachers (Boston: Allyn and Bacon, Inc., 1965), p. 19.

requirements for student teaching and all were training to be teachers) it was assumed these characteristics would be of no significant effect. The effects of other background factors such as religion, race, or socio-economic status, however, were unknown and remained unspecified.

Since most research studies draw samples of subjects that are biased in varying degrees with regard to some background characteristics, some understanding of the kind or amount of the influence of these factors needed to be developed.

The most obvious of the background factors that could have influenced the results was the sex of the test subjects. The sex characteristic was shown to have significant influence on M.M.P.I. test behavior, for example. So pronounced was the effect of the sex characteristic, that Hathaway and McKinley[22] found it necessary to furnish separate norms for men and women on almost all of the basic clinical scales. Thus, the present study also attempted to identify perceptual modes of males and females to determine the effect of sex grouping on the results.

---

[22]S. R. Hathaway and J. C. McKinley, The Minnesota Multiphasic Personality Inventory Manual (rev. ed.; New York: The Psychological Corporation, 1951), pp. 14-16.

The literature also suggested the appropriateness of a third classification. Allen[23] reported that students who plan to teach different subjects scored differently on the California Psychological Inventory scales. Student teacher groups in English, for example, scored differently on the C.P.I. than those in social studies, physical education, or math and science, who also scored differently from each other. Allen's findings supported the conclusion that no single combination of characteristics defined all student teachers without regard to choice of teaching major. Thus, the present study also attempted to utilize a projective technique to identify perceptual modes of student teachers classified according to major teaching field.

> The term projective technique has been applied to a wide variety of data-gathering procedures ranging from such classic instruments as the Rorschach and the TAT to such less widely used procedures as the sentence completion, word association tests, and picture drawing. While the "correct" usage of the term may occasion theoretical argument, general usage labels "projective" those techniques that require a person to bring structure to an unstructured test stimulus without knowing how his response will be evaluated.[24]

---

[23]Thomas K. Allen, "Personality Patterns of Student Teachers Planning to Teach Different Subjects," Journal of the Student Personnel Association for Teacher Education, IV (December, 1965), 92.

[24]Getzels and Jackson, op. cit., p. 554.

The general theoretical framework within which all projective techniques are fashioned is that there is a relationship between an individual's structuring of an ambiguous stimulus and his personality. How an individual structures or organizes meaningless stimuli into meaningful perceptions reveals basic aspects of his personality. Inkblots are most appropriate as stimuli because they are relatively meaningless and unstructured and permit a wide variety of individual responses. The subject, for example, when asked what he sees in an inkblot, must react in a personal, individual and unlearned fashion since there are no correct answers. And although the characteristics of the blot may condition the scope of the subject's response somewhat, his perceptions are selected and organized in terms of his fundamental personality structure.

L. K. Frank wrote,

> The Rorschach method offers a procedure through which the individual is induced to reveal his 'private world' by telling what he 'sees' in the several cards upon which he may project his meanings, significance, and feelings, just because they are not socially standardized objects or situations to which he must give culturally prescribed responses. The Rorschach method is essentially a procedure for revealing the personality of the individual as an individual, as contrasted with rating or assessing him in terms of his likeness or conformity to special norms of action and speech. It is just because a subject is not aware of what he is telling and

> has no cultural norms behind which to hide himself that the Rorschach and other projective methods are so revealing.[25]

Although Frank wrote specifically about the Rorschach in the above quotation, his remarks could also apply to other inkblot tests or techinques.

If, in realistic situations, a person's relationships with people are unsatisfactory, he may be unable or unwilling to perceive human forms in an inkblot. He may see an animal, a biological specimen or a forest from a tower. If, in his response, the subject uses the whole blot, major parts of it, or only a small detail, whether he uses the color, what he actually sees or is unable to see, what determines his response--all are held to reveal some of his personality structure.

However, direct personality hypotheses from projective test responses can only be tentative. Klopfer wrote that,

> Direct behavioral parallels may not always exist . . . and may not be the most important information revealed. The behavior of an individual in the . . . situation differs in one very important respect from the same individual's reaction to any unfamiliar situation in life. In life situations, a person tends to behave in a more or less socially acceptable way. He has learned, if he is a 'normal'

---

[25]L. K. Frank, "Projective Methods for the Study of Personality," Journal of Psychology, VIII (1939), 389, quoted in Bruno Klopfer and Helen H. Davidson, The Rorschach Technique: An Introductory Manual (New York: Harcourt, Brace and World, Inc., 1962), pp. 14-15.

> individual, to control himself, to appear pleasant and kindly rather than aggressive and hostile . . . and to conceal from public view those traits that are not prized. Thus a person's outward, observable behavior may often not reveal his true feelings and attitudes.[26]

In a projective testing situation a person does not know what the examiner is looking for or the typical way to respond. He reacts in his own particular manner and in so doing unconsciously reveals himself, even that part of himself of which he is not entirely aware.

Therefore there is both an advantage and a disadvantage in projective testing. There may be only slight correspondence between actual personality and behavior and predictions of personality and behavior because actual behavior is partly determined by the culture, environment and learned responses. Thus, any specific predictions of classroom personality and behavior of a student teacher must allow for learned responses to the environment. Since an analysis of the teaching environment was not a part of this study, direct personality predictions were not made. However, if perceptual trends (e.g. successful student teachers see more humans than less successful student teachers) were related to successful student teaching, there would be sufficient

---

[26]Bruno Klopfer and Helen H. Davidson, The Rorschach Technique: An Introductory Manual (New York: Harcourt, Brace and World, Inc., 1962), p. 15.

factual data to relate perceptual trends to observed classroom behavior.

Importance of the Study

The results of the study could have a significant impact on pre-training selection policies. In particular, if the null hypothesis (pages 14-15) tested were rejected, there would be cause for review of present personality assessment techniques. Certainly there would be cause to view teacher personality in a new light.

In any case, it is not being argued that measures of perception could become the answer everyone is seeking, but only that they might provide one more significant piece of information for use by administrators of teacher education programs.

Definition of Terms

1. Student teaching--that period of undergraduate guided teaching experience provided by Michigan State University as part of its teacher education program.

2. Student teachers--those students enrolled in Education 436 (Student Teaching) at Michigan State University during the Winter Term of the 1967-68 academic year.

3. University coordinators--those staff members employed by Michigan State University who coordinated

the programs and supervised students enrolled in student teaching.

4. Student teaching center--one of the seventeen geographical clusters of school districts in Michigan where Michigan State University placed student teachers.

5. Group Holtzman Inkblot variable--one of the seventeen variables measured by the group Holtzman Inkblot technique.

6. Rating of success--any one of the seven rating numbers assigned to a student teacher at the conclusion of his student teaching experience by his university coordinator. Success rating numbers are operationally defined in Appendix B.

Hypotheses

The following hypotheses were made concerning the relationship of computer scored group Holtzman Inkblot variables and the sample of student teachers.

1. There will be no significant differences among the mean scores on each of the following Holtzman Inkblot variables for student teachers grouped according to rated success in student teaching by the University coordinators.

A. (L) Location

B. (R) Rejection

C. (FD) Form Definiteness

D. (C) Color

E. (Sh) Shading

F. (M) Movement

G. (I) Integration

H. (H) Human

I. (A) Animal

J. (At) Anatomy

K. (Sx) Sex

L. (Ab) Abstract

M. (Ax) Anxiety

N. (Hs) Hostility

O. (Br) Barrier

P. (Pn) Penetration

Q. (P) Popular

2. There will be no significant differences among the mean scores on each of the above Holtzman variables for the male student teacher group and the female student teacher group.

3. There will be no significant differences among the mean scores on each of the above Holtzman variables for student teachers in various major teaching fields.

Assumptions

1. It was assumed that personality is an important variable in successful teaching.
2. It was assumed that perception is a function of personality.

3. It was assumed that the group Holtzman Inkblot variables measure perception.
4. It was assumed that the University coordinator's ratings of success on the basis of performance discriminated between student teachers who were more successful than others in their student teaching experience.

Limitations

As designed and conducted, the framework of this study possessed several recognized limitations. They were:

1. The study was limited in that it was not within the scope of the research to predict probable success as a teacher on-the-job. The study concerned the success of the student teacher in student teaching.
2. The study was also limited in that direct personality factors, as such, cannot be predicted without specifying environmental situations.
3. Finally, the study was limited in that no attempt was made to generalize to other populations.

Procedures for Collecting Data

The data for this study were collected in the following manner:

1. The group Holtzman Inkblot Test (HIT) was administered during a regular seminar meeting of the student teachers at each student teaching center. Each University coordinator distributed a letter (Appendix C) to his student teachers prior to the testing date and requested all the student teachers to be present. Nineteen separate administrations were given in the first weeks of the student teaching experience. In each administration, the test manual's directions to the examiner were closely followed. The screen projection method was used at each administration of the instrument. Unusual or adverse testing conditions were minimized and the administrations took place under normal conditions.

2. After the Holtzman records were collected using a specially modified answer sheet (Appendix D), the information was keypunched and verified by comparing the data card with the original response (Appendix E), loaded onto magnetic tape and computer scored. Computer output included a listing and a summary card of the Holtzman variable scores for each subject in the sample (Appendix E).

3. The rating of success of each student teacher was collected at the end of the term from each University coordinator who supervised any student teacher (Appendix G) and the success rating number entered on the summary card.

4. The final summary cards were then submitted for statistical analysis.

Description of the Sample

The Holtzman Inkblot Test was administered to each student enrolled in student teaching during the Winter Term of 1968. Of the 814 subjects tested, 12 did not receive terminal grades in student teaching and were dropped from the study. Useable Holtzman protocols were obtained from the remaining 802 subjects.

These 802 student teachers were fulfilling the requirements of student teaching leading to certification. Further description of the student teachers who served as the sample in this study is presented in Table 1.

The Instrument

The Holtzman Inkblot Test is an instrument designed to study personality through perception. The HIT group form was selected as the instrument for measuring perceptual trends because of an available computer scoring system which allows efficient and objective scoring of a large number of test protocols. Computer scored HIT protocols yield scores on perceptual modes described in Appendix A.

The group Holtzman Inkblot Test consisted of distributing the HIT RESPONSE SHEET (Appendix D) to each subject, directing the subjects to complete the

TABLE 1.--Distribution of sample population by teaching field and sex.

Teaching Field	Males	Females	Teaching Field	Males	Females
Elementary Grades			Secondary Subjects		
K-1-2	1	112	Instrumental Music	7	2
3-4	1	70	Vocal Music	1	1
5-6	3	48	Boy's Physical Education	20	0
7-8	6	35	Girl's Physical Education	0	17
Speech Correction	1	23	Speech	3	13
Secondary Subjects			Speech Correction	0	7
Distributive Education	4	1	French	1	17
Agriculture	15	0	German	0	1
Art	6	30	Spanish	2	10
Biology	15	10	Russian	2	2
Business (Secretarial)	0	10	Latin	1	1
Chemistry/Physics	4	1	Geography	0	2
English	12	59			
Social Studies/English Combination	3	7	Special Education		
General Science	2	3	Blind/Partially Sighted	0	4
History/Social Studies	25	52	Crippled/Homebound	1	7
Home Economics	0	45	Deaf/Hard of Hearing	0	2
Industrial Arts	13	0	Emotionally Disturbed	0	4
Beginning Mathematics	6	10	Mentally Retarded	3	17
Advanced Mathematics	13	8			
Sub Totals:	130	524	Sub Totals:	41	107

Males N = 171 Females N = 631 Total N = 802

identifying information, and projecting 45 inkblot slides, with an exposure time of 60 seconds each, on a screen. The subjects were requested to write on the HIT RESPONSE SHEET a brief, one sentence description of what they saw in each inkblot.

After the RESPONSE SHEETS were distributed and the identifying information completed, the subjects were introduced to the technique by the examiner who read the following directions prior to each group administration:

> You will be shown a series of inkblots, each of which will be projected on the screen for one minute. On the answer sheet, print or write clearly what each inkblot looks like to you. None of the inkblots have been deliberately drawn to look like anything in particular. Different people see different things in the blots; so there are no right or wrong answers.
>
> On many of the inkblots, you may see more than one thing. However, we want you to describe only one thing for each blot. Different aspects of the blot are important in what you see; for example, sometimes the shape seems most important, while on other blots, the color, shading, movement or some combination of these seems important.
>
> After writing your answer, circle 1, 1/2, or 1/4, in the box at the left to indicate how much of each inkblot you have used. If you used the whole inkblot, circle 1; if you used about one-half of the inkblot, circle 1/2; and if you used any part smaller than one-half of the blot, circle 1/4.
>
> Now let's look at some examples. (Blot X is projected on the screen in front of the group.) Here are three answers which might be given: A flying black bat. A pool of spilled oil. A steer's head. (Point to the areas of the blot used in each response - pause - then show blot Y.) Here are some answers people have given to this inkblot: A human figure (down through the center). A skeleton (bottom center portion). Several spots

> of red blood (all of the red parts). Funny little dwarfs with high hats (upper figures in black).
>
> Remember that you are to describe only one thing for each blot and that you are to circle the 1, 1/2, 1/4 in the box at the left to show how much of each inkblot you have used. Are there any questions?[27]

Any responses to questions were limited to clarifying the instructions. The examiner then exposed each of the forty-five slides for sixty seconds each. While the administration was in progress, the examiner monitored the group so that the subjects did not forget to indicate how much of each inkblot was used. One or twice it was desirable to say, "be sure to circle 1, 1/2, or 1/4 within the box to show whether you used the whole inkblot, or less than one-half."

The Compter Scoring System

In January, 1968, a computer scoring system for the HIT was made available to the public as a service by the Institute for Behavioral Research. This examiner, however, is grateful to the Institute for having access to the system before public announcement was made.

---

[27]Donald R. Gorham, Edward C. Moseley, and Wayne H. Holtzman, "Norms for Computer-Scored Holtzman Inkblot Technique," Perceptual and Motor Skills, XXVI (June, 1968), 1281.

The steps through which the computer scoring system progressed were presented by Gorham[28] and may be briefly stated as follows:

(1) An empirical list of words used by subjects was generated, based on over 500 subjects, with the assistance of a computer program.

(2) Scoring weights were assigned to each word to score seventeen HIT variables presented in Appendix A.

(3) The weights were adjusted to maximize the correspondence with an expert hand scorer. (The current English Scoring Dictionary has gone through five revisions.)[29]

(4) Pattern scoring computer sub-routines were created for certain variables which are independent, e.g., Color and Shading are dependent on FD. Form Definiteness and Location were appropriately corrected for Rejection. Integration was corrected by a multiple-regression equation to supplement the comparatively unsatisfactory scores achieved by the weighted dictionary alone.

The final computer program accepted the actual responses of each subject, either from IBM punched cards (Appendix E) or from data loaded on magnetic tape. Identifying information provided a two digit numeric code

---

[28]Donald R. Gorham, "Validity and Reliability of a Computer Based Scoring System for Inkblot Responses," Journal of Consulting Psychology, XXXI (February, 1967), 65-70.

[29]The original weights were assigned by Evelyn Hill, then a psychology trainee at Perry Point Veterans Administration Hospital. These weights were reviewed by Carol Hampton Swartz, an expert scorer at the HIT laboratory at the University of Texas. Seymour Fisher and Sidney Cleveland reviewed and re-weighted the Barrier and Penetration scores.

for each teaching field and a three digit numeric code for subject numbers. The sex of the subjects and the HIT form were also entered and were printed out with the scores for each subject (Appendix E).

Reliability and Validity of the Computer Scoring System

Gorham[30] presented extensive data concerning the reliability and validity of the computer scoring method. It will suffice here to set forth briefly the salient facts.

In order to test the reliability of the computer method, the records of 145 college students were hand scored by an experienced scorer on Holtzman's research staff. The correlations between this expert and computer scores are presented in Table 2.

For these seventeen variables it is evident that the correlations were high enough to demonstrate reasonable equivalence of the two methods of scoring. Eleven of the variables correlate favorably with hand scoring values although for the remaining six variables, some computer scores were too high and some too low when compared with the hand scorer. As Gorham suggested,[31] however, such an approach presumes infallibility of the hand

---

[30] Gorham, op. cit.

[31] Ibid., p. 67.

TABLE 2.--A comparison of hand scoring and computer scoring for 17 HIT variables.[32]

Variable	r
L	.97
R	1.00
FD	.84
C	.79
Sh	.58
M	.93
I	.67
H	.89
A	.90
At	.80
Sx	.94
Ab	.54
Ax	.79
Hs	.65
Br	.63
Pn	.62
P	.77

scorer. To secure a broader base, a sub-sample of the 145 college students was selected and their protocols scored independently by three hand scorers and the computer. Table 3 indicates that for nine variables the correlation of computer scores with three hand scorers was equal to or greater than the average inter-scorer correlation. For five variables, however, the computer was only a few points lower and for the remaining two variables, the computer was least efficient.

A further odd-even reliability test was performed on the basic 145 college student sample and the results

[32] Ibid., p. 66.

are presented in Table 4. The findings demonstrated that the computer matches hand scoring in terms of split-half reliability and further demonstrated that the variable scores which are the most difficult to simulate by computer were the ones for which hand-scorer reliability is lowest, especially shading.

TABLE 3.--A comparison of computer scoring with three hand scorers.[33]

Correlations with Computer of 3 Hand Scorers					Correlations Between 3 Hand Scorers			
C & 1	C & 2	C & 3	C & AVG.	Variable	1 & 2	1 & 3	2 & 3	AVG.
.99	.93	.95	.97	L	.94	.96	.94	.95
1.00	1.00	1.00	1.00	R	1.00	.99	.99	.99
.81	.73	.76	.81	FD	.78	.87	.86	.84
.77	.87	.79	.89	C	.74	.69	.82	.75
.59	.50	.63	.70	Sh	.14	.50	.49	.38
.92	.75	.90	.89	M	.81	.96	.79	.85
.75	.58	.51	.69	I	.56	.51	.63	.57
.92	.90	.89	.92	H	.95	.96	.94	.95
.90	.87	.90	.91	A	.94	.94	.91	.93
.89	.79	.75	.87	At	.82	.86	.75	.81
.99	.99	.98	.99	Sx	1.00	1.00	.99	1.00
.60	.60	.61	.66	Ab	.73	.88	.71	.77
.80	.72	.69	.81	Ax	.73	.73	.73	.73
.64	.57	.46	.59	Hs	.79	.83	.79	.80
.64	.73	.72	.81	Br	.61	.55	.70	.62
.56	.35	.63	.58	Pn	.62	.60	.62	.62

Thus, these studies demonstrated that computer scored HIT records produce essentially the same results as expected from individually scored records.

[33] Ibid.

TABLE 4.--Odd-even reliability coefficients for hand and computer scoring.[34]

Computer Scored	Variable	Hand Scored
r		r
.88	L	.88
.87	R	.76
.72	FD	.77
.61	C	.60
.38	Sh	.28
.64	M	.70
.72	I	.43
.70	H	.75
.63	A	.64
.45	At	.52
.85	Sx	.94
.59	Ab	.76
.41	Ax	.49
.35	Hs	.47
.48	Br	.39
.39	Pn	.45
.40	P	.25

Rating of Success

A survey of the literature revealed that methods of teacher evaluation vary considerably and present a major obstacle in teacher education studies. Some researchers use extensive check lists while others prefer essay and other unstructured forms of evaluations. Ryans[35] and

[34] Ibid., p. 68.

[35] David G. Ryans, "The Criteria of Teaching Effectiveness," Journal of Educational Research, XLII (May, 1949), 690-699.

Fattu,[36] in reviewing the literature on teacher effectiveness, found that teacher rating devices were the most frequently used means of assessing teacher behavior and efficiency.

Ryans[37] suggested that there are two general approaches to the assessment problem in teaching. The first is to observe the teacher; the second, to observe the teacher's effect on pupils. And of these, ratings of behavior through observation of the teacher are most frequently used. Ryans further suggested, though, that ratings of observed behavior should be made by trained and experienced observers to be effective.

Another possible measure of potential teacher success is the rated performance of student teachers during student teaching. Research[38] suggested that success as a student teacher usually leads to success as a teacher on-the-job.

Yet, it is true that there is no universally accepted measure of teacher effectiveness and investigation has not yielded meaningful, measurable criteria of teacher

---

[36]Nicholas A. Fattu, "Effectiveness - An Elusive Quality," The Education Digest, XXVII (January, 1962), 24-26.

[37]Ryans, op. cit.

[38]Morsh and Wilder, op. cit., p. 74.

effectiveness. Osmon[39] wrote, however, that perhaps instead of trying to develop foolproof schemes to rate teachers, major considerations for determining effectiveness should be given to the teacher's actual performance.

The student teachers in this study were evaluated on their performance in student teaching by competent, professional supervisors, using the SUCCESS RATING CHART (Appendix B).

The chart was developed by the author and used by Underhill[40] and Gustafson[41] in their unpublished research studies. The basis for the instrument (the criteria of success listed on the chart) was originally conceived, developed and modified by the Michigan State University coordinators. Their familiarity with the instrument and the implications of each criterion to rate the student teachers insured greater reliability. Face validity, however, was assumed.

---

[39]Robert V. Osmon, The Improvement of Secondary Teaching (St. Louis: Education Publishers, 1962).

[40]Robert G. Underhill, "The Relation of Elementary Student Teacher Empathy (Affective Sensitivity) Change to Supervising Teacher Empathy and Student Teaching Success" (unpublished Ed.D. dissertation, Michigan State University, 1968), p. 58.

[41]Kent Gustafson, "Simulation of Anxiety Situations and Its Resultant Effect on Anxiety and Classroom Interaction of Student Teachers" (unpublished Ph.D. dissertation, Michigan State University, 1969), p. 66.

The method of rating the student teachers consisted of the University coordinators rating the student teachers they supervised using the SUCCESS RATING CHART (Appendix B), and the criteria of success indicated on the chart as a guide to assign one of the seven success rating numbers, which, in the coordinator's best judgment, represented the best description of the student's performance in student teaching. Table 5 presents a summary of the distribution of success rating numbers for the sample.

TABLE 5.--Distribution of success rating numbers.

Number	Description	N
1	One of the very best student teachers I have ever seen.	51
2	Highly Successful	232
3	Successful	432
4	Less Successful	74
5	Passed but should not be in teaching	13
6	Failed	0
7	Grade Postponed	12
	TOTAL	814

However, student teachers were also guided in their learning experiences by supervising teachers who are competent and professional supervisors. Nevertheless,

supervising teachers' rating of success would be limited in that the teachers were exposed to only one student teacher at a time and generally made no observations of other student teachers. Hence, supervising teachers have little observational base on which to compare student teachers and rate their relative success. The University coordinators, however, were in close communication with the supervising teachers and the coordinators supervised the student teaching of all students assigned to him. Thus, success ratings from supervising teachers were not solicited; only success ratings from those in a position to compare the performance of student teachers were obtained.

CHAPTER II

REVIEW OF RELATED LITERATURE

The literature devoted to studying the characteristics of the effective teacher is vast and voluminous. The question of the characteristics of the successful teacher has probably brought about more research than any other in education. Gage[1] noted, for example, that not only is the literature on teacher competency overwhelming, even bibliographies on the subject are becoming unmanageable. Thousands of investigations are reported dealing with characteristics of teachers, effects of teaching, goals of education and other related problems. Yet few, if any, facts are established concerning teacher success.

> Despite the critical importance of the problem and a half-century of prodigious research effort, very little is known for certain about the nature and measurement of teacher personality, or about the relation between teacher personality and teaching effectiveness.[2]

It appears that the multidimensionality of the research stems from a number of serious obstacles which

---

[1]N. L. Gage, Address appearing in "Proceedings," Research Resume (Burlingame, California: California Teachers Association, 1960), p. 16.

[2]Getzels and Jackson, op. cit., p. 574.

face the research worker in this area: (1) the problem of definition of personality, (2) the lack of a precise and measurable definition of teacher success, and (3) the problem of subjectivity in evaluating the teacher's performance once the criteria of success are specified. And all of the difficulties are cumulative in predictor investigations of successful teacher personality characteristics.

No widely accepted measure of teacher competence is recognized and no methods of predicting teacher success are generally accepted. Since the topic is complex and encompasses many theoretical and methodological problems, the researchers utilize an array of approaches extending from general global studies to sophisticated, technical inquiries. The research focal points, however, may be classified into two broad areas: (1) rating and criteria studies, and (2) predictor investigations. And each broad classification includes both general and specific inquiries on characteristics, needs, adjustments, values, personality traits and related variables of the effective teacher.

Problems of Personality Definition

To some personality theorists, environment is highly significant in defining personality. To others, early life experiences are most important. Some treat

personality within relative stable dimensions of habits, learned responses and traits, while others consider dynamic approaches such as needs and drives. Some emphasize conscious aspects of behavior over the unconscious. Abt[3] wrote that, in general, personality is becoming increasingly conceptualized as a gestaltic process influenced by both environmental interaction and the intensity of need drives.

The concept of personality, as distinguished from other areas of psychology, appears to involve at least two basic ideas: uniqueness and integration. Allport encompassed these two ideas when he defined personality as, ". . . the dynamic organization within the individual of those psychological systems that determine his characteristic thought and behavior."[4]

Since psychology is rapidly achieving importance in applied endeavors, psychologists seek to assess personality in a precise and demonstrable manner and a considerable amount of effort consists of personality assessment for evaluation, selection and vocational guidance activities.

---

[3]Lawrence E. Abt and Leopold Bellack, Projective Psychology (New York: Alfred A. Knopf, 1950), pp. 59-63.

[4]Gordon W. Allport, Pattern and Growth in Personality (New York: Holt, Rinehart and Winston, 1961), p. 28.

It seems, too, that almost everyone in teacher education agrees that personality is somehow linked to successful teaching.[5] If one could determine the precise characteristics of the successful teacher, there would be a base for evaluation, selection, and guidance of teacher candidates.

However, two major difficulties appear once personality is defined. How does one judge what is successful teaching (rating and criteria problems)? And, how can one predict successful teaching (predictor studies)?

Criteria of Teacher Effectiveness Problems

> By common definition a criterion is any standard used for judging. For the scientist, however, such a definition is inadequate. A criterion which is to be used for scientific judgments cannot be just any standard. It should be the best possible standard for the particular class of judgments that are to be made. This means that the scientist must be able to justify his choice of a criterion by demonstrating its logical relevance to the problem at hand and by showing that it possesses measurement characteristics which are technically adequate.[6]

How does one then, determine what is effective teaching? Some researchers, perhaps unwisely, assumed that everyone knows what describes a successful teacher. The evidence is prima facie. Everyone says he is. Opinions

---

[5]Bruce J. Biddle and William J. Ellena, Contemporary Research on Teacher Effectiveness (New York: Holt, Rinehart and Winston, 1964), p. 9.

[6]Morsh and Wilder, op. cit., p. 7.

have been offered by just about everyone capable of offering an opinion: pupils, teachers, principals, superintendents, parents, professors of education, teacher agency executives and custodians.

For over fifty years researchers attempted to identify the personality characteristics of successful teachers by making lists of traits based entirely on opinions.[7] In most cases, these lists were subjectively estimated characteristics of such a vague, general nature as to render any precise measurement of them impossible. For example, it is said after the usual inventory tabulation that good teachers are friendly, cheerful, sympathetic, intelligent and morally virtuous rather than the opposite. "What is needed, however, is not research leading to the reiteration of the self-evident, but to the discovery of specific and distinctive features of teacher personality."[8]

Thus, the trend in present day research in teacher characteristics is directed away from opinion studies and toward the use of psychological theory in the development of systematic sets of hypotheses to be tested with objective tests and observational techniques.

Teacher educators agree that identifying the effective instructor is crucial to teacher education. This

---

[7] Ibid., pp. 108-113.

[8] Getzels and Jackson, op. cit., p. 574.

identification could be used for many purposes: curriculum construction in teacher education, certification, tenure, selection and promotion. Steeves[9] wrote that for the last twenty meetings of the Association for Student Teaching, more than half of the programs consider the issue of the criteria of teacher effectiveness. Yet, after more than fifty years of research, there exist no standards which are commonly agreed upon as the criteria of teacher success.

But the problem is not unique to teacher education. Concepts of concurrent or predictive validity of any selection procedure imply a satisfactory criterion measurement. Stevens pointed out that concepts such as criterion measurement are defined in terms of the operations that produce them. "Thus, teacher effectiveness has no meaning apart from the criterion measures of success as a teacher."[10] These measures, Stevens further suggested, should possess four basic attributes: (1) relevance, (2) reliability, (3) freedom from bias and (4) practicality. And of these, it seems that relevance is the most important attribute.

---

[9]Frank L. Steeves, "Crucial Issues in Student Teaching," Journal of Teacher Education, XVI (September, 1965), 307-310.

[10]S. S. Stevens, "Psychology and the Science of Science," Psychological Bulletin, XXXVI (April, 1939), 221-263.

Relevance of a criterion implies that it reflects the behaviors required in the achievement of objectives. However, there is lack of agreement among educators on a hierarchy of goals and objectives for teaching. Some teachers consider behavioral changes in attitudes to be most important. Others believe intellectual skills to be of primary importance. Some hold that personal and social adjustment is the first goal of education while some maintain that cultural appreciations are the ultimate objectives of education.

> Should a teacher's tasks be defined in terms of the ultimate goals of education or in terms of a specific effect upon the pupil? Should a teacher be called upon for equal competence with all types of pupils, or should special competencies be allowed in working with the underprivileged, the handicapped, the exceptional pupil? Are teachers to be judged in terms of the aims of 'classical' or 'progressive' education? Are teachers to be judged for competence against each of these simultaneously?[11]

The proliferation of teaching objectives and the wide general acceptance of none has made the task of selecting relevant teacher effectiveness criteria exceedingly difficult. Regardless of the classification or methodology of criteria studies, however, all involve some commitment on the following issues: is teacher effectiveness multidimensional or unidimensional; should teaching effectivensss be evaluated primarily against cognitive goals or affective goals? The issues are

---

[11]Biddle and Ellena, op. cit., p. 3.

unresolved. Nevertheless, researchers press forward realizing that any proposed criteria variable is a partial variable.

> Relevance for a particular variable should be defended logically without undue apology for its lack of comprehensiveness since it seems unlikely that we shall soon find a generally accepted formula for weighing school goals according to their importance for overall effectiveness.[12]

Criteria of teacher effectiveness studies may be classified by any one of several proposed schemes. Mitzel[13] reviewed the work of Remmers and others who suggested that criteria can be classified into ultimate or proximate while Thorndyke proposed similar classification into immediate, intermediate and ultimate. Kelly and Fiske called criteria first order and second order with the additional classification of either specific or general. Mitzel[14] also reviewed the classification proposed by Brownell who suggested that criteria be classified according to product criteria, process criteria or presage criteria.

A possible product criterion (a set of goals toward which student teaching is directed) of potential teacher success is the performance of student teachers during the time of their practical laboratory experience. A wealth

---

[12]Mitzel, op. cit., p. 1482.

[13]Ibid.

[14]Ibid.

of research[15] indicates that success as a student teacher generally leads to success as a professional teacher. Yet, judgment of success of a student teacher encounters the same criterion difficulties as judging success of a teacher.

The student teaching staff at Michigan State University has developed a set of product criterion variables for judging student teaching success that considers the overall impact of the following factors:

A. Working With People

B. Establishing Classroom Climate

C. Planning Instruction

D. Managing Instruction

E. Command of Subject and Teaching Materials

F. Personal Qualities

G. Professional Qualities

H. General Effectiveness as a Teacher

Hicks and Blackington[16] discussed each product criterion in some detail and emphasized that a student teacher who works well with people, for example, was one who established satisfactory rapport with pupils, staff and parents. And this rapport, the writers suggested,

---

[15]Morsh and Wilder, op. cit., p. 74.

[16]William Vernon Hicks and Frank H. Blackington, Introduction to Education (Columbus, Ohio: Charles E. Merrill Books, Inc., 1965), pp. 38-56.

related closely to the student teacher's ability to establish classroom climate and manage instruction.

These criteria of success were used by LePere and Cox[17] in a study of pre-service training of elementary teachers. The variables form the basis of the Confidence Level Inventory for Teachers which the researchers administered as a pre-student teaching and post-student teaching test to determine changes in self-rating relative to variations in undergraduate methods courses. When the student teachers rated themselves on questions related to each variable, correlations between pre- and post-student teaching evaluations of control and experimental groups ranged from +.23 for Working with People to +.78 for Planning Instruction.

In summary, it appears that the most appropriate criteria of teacher effectiveness are those that have relevance to significant educational objectives, and since the profession appears to have a proliferation of these objectives, the specification of clear, concise and objectively measurable criteria variables is exceedingly difficult. Biddle and Ellena wrote that, "until effects

---

[17]Jean M. LePere and Richard C. Cox, Training Elementary Teachers: Comparisons of Separate and Block Methods Courses (East Lansing, Michigan: Bureau of Educational Research Services, Michigan State University, 1964).

desired of the teacher are decided upon, no adequate definition of teacher comptence is possible."[18]

Problems of Rating Teacher Effectiveness

A researcher could raise several serious objections to any criterion used to assess teaching success. Perhaps the main objection is that whatever the criterion may be, the method of assessing the criterion is subjective because assessment depends upon the assessor's interpretation of the criterion and the impression made on the assessor by the teacher. Different assessors may receive different impressions. Different assessors may have different interpretations of the criterion.

In their comprehensive review of assessment research literature, Morsh and Wilder reported seven primary methods of assessing success once the criterion has been specified: (1) rating devices, (2) administrative ratings, (3) peer ratings, (4) student ratings, (5) self-ratings, (6) systematic observations and (7) student gains evaluation. Morsh and Wilder commented on each method:

> Rating devices have failed to provide means for identifying significant items in setting up instructor rating devices. There is no general agreement as to what constitutes the essential characteristics of a competent teacher.

---

[18]Biddle and Ellena, op. cit., p. 4.

Administrative ratings . . . are the most widely used measure of teacher competence . . . and show . . . that teachers can be reliably rated by administrative . . . personnel (usually with r's of .70 or above). However, ratings . . . are apt to be contaminated by halo.

Peer ratings have been little used . . . and they are probably not too useful since teachers . . . have misgivings about passing judgments on fellow teachers . . . halo influences peer ratings just as it does administrative ratings.

The use of student ratings . . . appears to be growing. However, considerable halo effect is usually found depending whether or not grades are given. There is little research evidence that student ratings could be used to improve supervisors ratings.

Self-ratings. While there is some tendency for instructors to over-rate themselves, self-ratings show negligible relationships with administrative ratings, student ratings, or measures of student gains. On the basis of the few available studies of self-ratings of instructors, the obvious, undisguised self-rating would seem to offer little encouragement for evaluation or research purposes.

Systematic observations. No single specific observable teacher act has yet been found whose frequency of occurance is invariably significantly correlated with student gains. Most of the observations made have been dependent upon the subjective judgment of the observer.

Student gains. The great discrepencies in the findings of investigators who have examined the student gains criteria emphasize the extreme variability of relationship with other criteria to indicate instructor ability within the limits of measures so far used. The relationship between administrative opinion of an instructor's competence and the amount of subject matter that the instructor will impart to his students cannot be predicted.[19]

---

[19]Morsh and Wilder, op. cit., pp. 3-4.

Evans surveyed the methods of assessing teacher ability and concluded that of the possible methods, "none . . . is entirely satisfactory . . . the opinion of experts is at present probably the most suitable for general use."[20]

Barr[21] also reviewed the extensive literature on teacher effectiveness assessment and concluded that the results are disappointing. Perhaps, Barr suggested, teaching is too complex for analysis and the lack of objective success criteria compounds the difficulty. Other recent findings tend to confirm Barr's conclusion and Evans summed up the literature when she wrote, "It may well be that no general assessment of teaching ability is possible."[22]

Thus it appears that teacher effectiveness research is handicapped by the lack of success criteria, rating problems, and nonagreement on the objectives of education. Combs wrote that, "the very failure of research to define common factors is, itself, a demonstration that a good teacher is primarily a personality."[23]

---

[20]Kathleen M. Evans, "A Critical Survey of Methods of Assessing Teacher Ability," British Journal of Educational Psychology, XXI (June, 1951), 94.

[21]Arvil S. Barr, "The Measurement and Prediction of Teacher Efficiency, A Summary of Investigations," Journal of Experimental Education, XVI (June, 1948), 203-283.

[22]Evans, op. cit., p. 94.

[23]Combs, op. cit., p. 372.

However, several writers feel that, subjective handicaps notwithstanding, research on assessing teacher effectiveness can be undertaken. Administrative or supervisory ratings seem to offer a method which is somewhat more reliable and practical than others. Rolfe[24] found that teacher effectiveness rating scales used by competent and experienced supervisors show a significant correlation within teaching ability, and Hampton[25] reported that it is possible to develop rating scales which contain relatively few items and yet accurately measure teacher effectiveness.

> If one is concerned with teacher effectiveness as the changes brought about by the teacher in the teacher's own pupils, then ratings are less relevant than either measures of student change or controlled observations of student behavior. Ratings are someone's estimate of the effects on students of those teacher characteristics the rater happened to observe, and which he deemed important. Without demonstration that these estimates have relationship to student achievement, they cannot really be considered as satisfactory substitutes for measures of pupil change. On the other hand, if one is considering that part of teacher effectiveness which the teacher contributes to the growth of all pupils by participation in the efforts of the educational group, then ratings would seem to be seomwhat more relevant. In this latter case, the influence of the teacher is a function of the quality of the teacher's relations with students

---

[24] J. F. Rolfe, "The Measurement of Teaching Ability," Journal of Experimental Education, XIV (September, 1945), 52-74.

[25] N. D. Hampton, "An Analysis of Supervisory Ratings of Elementary Teachers Graduated from Iowa State Teachers College," Journal of Experimental Education, XX (December, 1951), 180-215.

> in general, with other teachers, supervisors and the community. Differential effectiveness is a matter of differential contribution to the over-all goals of the school or educational system. Since such contribution is almost inevitably in a cooperative setting, and since its effects are diffuse and (almost certainly) unmeasurable, there would appear to be logical justification for an attempt to get estimates of effectiveness in this area by the use of ratings obtained from others in the educational situation.[26]

The usual method of assessing teaching success, Evans[27] wrote, is the judgment of experts who observe the work of the teacher. Evans also reported the conclusion of a study by Sandiford and Trump "that the judgment of experts is the only valid criterion of teaching success presently available."[28]

Thus, a rating scale appears to have usefulness although, like any other method, it is subject to many limitations. Pettinger wrote:

> The proper function of a score card for teacher measurement is not to substitute such a formula for a supervisors judgment, but to . . . aid in discovering and assembling all the data upon which intelligent judgment should be based.[29]

As previously discussed, a possible criterion of potential teacher success is the performance of student

---

[26]Morsh and Wilder, op. cit., p. 9.

[27]Evans, op. cit.

[28]Ibid.

[29]B. F. Pettinger, "Problems of Teacher Measurement," Journal of Educational Psychology, VIII (January, 1917), 110.

teachers during the time of their practical laboratory experience. Research indicates that assessments made by competent supervisors appears to be a practical method of evaluating effectiveness and since student teaching success usually leads to success as a professional teacher,[30] the present study utilized professional judgments of success on the SUCCESS RATING CHART as the rating method.

The SUCCESS RATING CHART (Appendix B) considered the overall impact of the designated criteria (p. 39) and requested the University Coordinator to assign a success rating number (Appendix B) to each student teacher. This scale was developed specifically for this study and has been used in interim investigations by Underhill[31] and Gustafson.[32]

Problems in the Prediction of Teacher Effectiveness

Biddle and Ellena, in reporting the efforts of contemporary research on teacher effectiveness, suggest that "the problem of teacher effectiveness is so complex that no one today knows what the competent teacher is."[33] Hamacheck, however, wrote that, ". . . we do know what

---

[30]Morsh and Wilder, op. cit., p. 74.

[31]Underhill, op. cit., p. 58.

[32]Gustafson, op. cit., p. 66.

[33]Biddle and Ellena, op. cit., p. 2.

the competent teacher is . . ."[34] Competent teachers, Hamacheck pointed out, appear to have positive views of others, are not critical or suspicious of the motives of others, generally adhere to democratic classroom procedures and have the ability to see things from the other person's point of view. They see students as capable of doing for themselves once the students feel trusted, respected and valued. Hamacheck concluded that, "a good teacher is a good person, a good teacher is flexible."[35]

Symonds[36] reported that good teachers like children, possess good integrated personality organization, are personally secure and self-assured; inferior teachers are the opposite. Kemp[37] wrote that good teachers are dominent, tolerant and self-confident while poor teachers tend to be over-critical, socially passive and uncooperative. Witty[38] asked pupils to indicate the characteristics of

---

[34]Don Hamacheck, "Characteristics of Good Teachers and Implications for Teacher Education," Phi Delta Kappan, L: 6 (February, 1969), 341.

[35]Ibid., p. 343.

[36]Percival M. Symonds, "Characteristics of the Effective Teacher Based on Pupil Evaluations," Journal of Experimental Education, XXIII (June, 1955), 289-310.

[37]L. D. Kemp, "The Prediction of Teaching Success," The Forum of Education, VI (August, 1947), 14.

[38]Paul Witty, "Some Characteristics of the Effective Teachers," Educational Administration and Supervision, XXXVI (April, 1950), 193-208.

teachers who help them the most. The pupils indicated that these teachers are cooperative, kind, patient, fair, flexible, have wide interests, a sense of humor and use recognition and praise. Barr[39] extensively studied magazine articles on the teaching of social studies and summarized the personal factors relative to teacher effectiveness: subject knowledge, intelligence, skill in expression, social adjustment, emotional stability, leadership qualities, interest in teaching, health, energy, motivation and empathy. Stout[40] suggested that emotional maturity, moral and ethical fitness, intelligence, ability to work with children, motivation, and professional interests distinguish the good teacher.

Ryans wrote that,

> Measured intellectual abilities, achievement in college course, general cultural and special subject matter knowledge, emotional adjustment, attitudes favorable to students, generosity in appraisals of the behavior and motives of other persons, strong interest in reading and literary matters, interest in music and painting, participation in social and community affairs, early experience in caring for children and teaching (such as reading to children, taking a class for the teacher), history of teaching in family, size of school and size of community in which teaching,

---

[39]Arvil S. Barr, "Pupil Changes and Teacher Ability," Journal of Experimental Education, XIV (September, 1955), 52-74.

[40]Ruth A. Stout, "Practices for Selection in Teacher Education," Teacher Education: The Decade Ahead (Washington, D.C.: National Commission on Teacher Education and Professional Standards, National Education Association, 1955), p. 1490.

> cultural level of community, and participation in avocational activities, all appear to be characteristics of the teacher which are likely to be positively correlated or associated with teacher effectiveness in the abstract.[41]

In 1960, Ryans completed what Getzels and Jackson refer to as, "the single most extensive study of teachers to date."[42] The Teacher Characteristics Study[43] involved thousands of teachers in over seventeen hundred schools and took more than six years to complete. Ryans reported the ultimate purposes of the study:

> The Teacher Characteristics Study was conducted with two possible results in mind: first . . . identifying teachers who . . . have characteristics . . . deemed important and desirable . . . and, second, contribute to improved procedures for selecting teacher candidates and to the improvement of professional courses and curricula.[44]

Observer ratings distinguished three groups of teachers: those above, at, and below the mean on each of three central classroom behavior dimensions of the study (friendly versus aloof, systematic versus slipshod, stimulating versus dull). The more notable differences between the high and the low teachers are reported by Ryans:

---

[41]David G. Ryans, "Prediction of Teacher Effectiveness," in Encyclopedia of Educational Research, ed. by C. W. Harris (3rd ed.; New York: Macmillan Company, 1960), p. 1490.

[42]Getzels and Jackson, op. cit., p. 566.

[43]David C. Ryans, Characteristics of Teachers (Washington, D.C.: American Council on Education, 1960).

[44]Ibid., p. 11.

> There was a general tendency for high teachers to: be extremely generous in appraisals of the behavior and motives of other persons; possess strong interest in reading and literary affairs; be interested in music, painting, and the arts in general; participate in social groups; enjoy pupil relationships; prefer nondirective (permissive) classroom procedures; manifest superior verbal intelligence; and be superior with respect to emotional adjustment. On the other hand, low teachers tended generally to: be restrictive and critical in their appraisals of other persons, prefer activities which did not involve close personal contacts; express less favorable opinions of pupils; manifest less high verbal intelligence; show less satisfactory emotional adjustment; and represent older age groups.[45]

Thus it seems that much is known about the characteristics of the effective teacher and thousands of researchers have attempted to demonstrate these characteristics with objective tests.

The studies falling within the scope of the teacher's personality and characteristics are too numerous for individual mention, much less adequate treatment. However, Barr,[46] Domas and Tideman,[47] Morsh and Wilder,[48] and Getzels and Jackson[49] have compiled extensive annotated bibliographies on the topic. Morsh and

---

[45] Ibid., pp. 397-398.

[46] Barr, "The Measurement . . ," pp. 203-283.

[47] Domas and Tideman, op. cit., pp. 101-218.

[48] Morsh and Wilder, op. cit.

[49] Getzels and Jackson, op. cit.

Wilder commented on the research on several major predictors of teaching success:

1. Intelligence

> Whether or not intelligence is an important variable in the success of the teacher apparently depends upon the situation. In general, there appears to be only a slight relationship between intelligence and rated success of a teacher. Correlation coefficients for high school teachers tend to be somewhat higher and somewhat less variable than those reported for elementary teachers. This variable appears to be of little value as a single predictor of teacher effectiveness.

2. Education

Amount of education in semester hours or years appears to have a very slight relation to teacher effectiveness.

> Beyond certain more or less obvious knowledge requirements . . . amount of education . . . seems to be unimportant in discriminating between good and poor teachers.

3. Scholarship

> . . . Accurate prediction of success in practice teaching cannot be made on the basis of an individual's scholastic record in high school. Almost all studies report low positive correlation between measures of on-the-job performance of teachers and earlier scholarship in high school, or college, or in standing obtained in specific college courses. No investigator has shown that the attainment of a particular standing in high school or college or the mastery of a particular course or group of courses is essential to teaching competency. The positive correlations that are found tend to reflect the relationship of general intelligence to both academic and teaching success.

4. Age and Experience

It appears that a teacher's rated effectiveness increases at first rather rapidly with experience and then more slowly up to five years or beyond. Then there is a leveling off, and the teacher may show little change in rated performance for the next 15 or 20 years, after which, as in most occupations, there tends to be a decline.

5. Knowledge of Subject Matter

Whether or not knowledge of subject matter is related to teacher competence seems to be a function of the particular teaching situation. Some studies suggest that too much knowledge on the part of the teacher may result in teaching "over the heads" of students.

6. Professional Information

Scores on tests of professional information appear to bear some slight relationship to supervisory ratings or ranking of teacher competence. Contradictory results have been obtained, however, when such scores are correlated with pupil gain.

7. Extracurricular Activities

In general, investigators have found a low positive relationship between an individual's participation as a student in extra-curricular activities and his later teacher effectiveness.

8. General Culture

Studies reviewed appear to indicate that the relation of Cooperative General Culture test scores to teacher effectiveness differs little from those reported for other subject matter tests.

9. Socioeconomic Factors

Studies of the relationship of socioeconomic status to criteria of teacher effectiveness shows little correlation,

. . . unless it is that those from higher status groups have greater probability of success in life than those less fortunate.

10. Sex

No particular differences have been shown when the relative effectiveness of men and women has been compared.

11. Marital Status

There appears to be no evidence that married teachers are in any way inferior to unmarried teachers.

12. Teaching Aptitude

Results obtained from measures designed to predict teaching ability shows great disparity. Data thus far available either fail to establish the existence of any specific aptitude for teaching with any degree of certainty or indicate that tests used were inappropriate to its measurement.

13. Interest in Teaching

In most of the studies reviewed, interest in teaching was measured by interest test scores which indicated similarity of teachers and persons undergoing the interest test. Correlations resulting from the use of several standard interest tests either cluster around zero or are so inconsistent as to render such tests of rather doubtful value as predictors of teaching success. The common factors that were found through factor analysis to underlie the reasons given for choosing the teaching profession are perhaps provocative of further research but were based on too few cases to justify any clear cut interpretation.

14. Voice and Speech Characteristics

It appears that the quality of the teacher's voice is not considered too important by administrators, teachers or students.

15. The Photograph

Studies of the use of the photograph as a predictor of teacher effectiveness have failed to demonstrate that photographs have any predictive value.

16. Statistical Analyses of Teacher Abilities

Such instructor factors as empathy, professional maturity, general knowledge, mental ability, social adjustment and the like have been identified through factor analysis by various investigators. The statistical analyses thus far reported, however, suffer from inadequacies of criteria, testing instruments or the number of cases.

17. Opinion Studies of Teacher Personality Characteristics

The attempts made to identify characteristics of successful and unsuccessful teachers by making lists of traits based on opinion appear largely sterile in terms of usability for evaluation or selective purposes.

18. Causes of Teacher Failure

In most of the studies of unsuccessful teachers, poor maintenance of discipline and lack of cooperation tend to be found as the chief causes of failure. Health, educational background, training, age and knowledge of subject matter, on the other hand, appear to be relatively unimportant factors in terms of teacher failure.

19. Teaching Attitude

Attitude toward teachers and teaching seems to bear a small but positive relation to teacher success measured in terms of pupil gain.[50]

In these predictor studies reviewed by Morsh and Wilder, the criteria problems are handled with widely varying degrees of sophistication. Measures found

---

[50]Morsh and Wilder, op. cit., pp. 5-7.

acceptable as criteria of instructor effectiveness by one investigator are often considered as unvalidated by others. But there are several limitations too, Ryans[51] suggested, which apply to the predictability of all criterion measures and thus affect the usefulness of the findings:

> The predictability of teacher effectiveness undoubtedly is affected by the multidimensionality of the criterion. There is accumulating evidence that prediction can be accomplished with better than chance results for specified dimensions or components of the criterion. On the other hand, the prediction of overall teacher effectiveness is possible only to the extent that some general agreement can be reached regarding the dimensions comprising overall effectiveness and how they should be combined to form a composite.
>
> The predictability of teacher effectiveness varies depending on the degree of control it is possible to exert in dealing with the multiplicity of predictors and in the multidimensionality of the criterion.
>
> The predictability of the criterion varies with the kind of measure employed in obtaining the criterion data.
>
> The predictability of the criterion varies with the adequacy (reliability and validity) of measures of the criterion and the predictor variables.
>
> The predictability of the criterion is so limited by conditions associated with measurement of the criterion, measurement of predictors, and practical conditions, that relationships representing common variance of perhaps one fifth or one fourth of the total variance probably approach the maximum to be expected in chance instances.

---

[51]Ryans, Encyclopedia of Educational Research, p. 1490.

> The predictability of a dimension of the criterion of teacher effectiveness from a specified predictor probably varies depending upon the cultural millieu which provides the setting for an investigation, particularly the values and objectives prominent in the teacher-training curriculum at the time the teachers studied in college.
>
> The predictability of the criterion varies directly with the degree of similarity between the sample with respect to which predictors are derived and the sample to which the predictors are applied in attempting to determine predictor criterion relationships.
>
> The predictability of a criterion dimension varies with the particular teacher population being studied.
>
> The predictability of the criterion varies inversely with the time interval separating the obtaining of the predictor measurements and criterion measurements.
>
> The predictability of the criterion probably varies depending upon the association of incentive or non-incentive conditions with the obtaining of predictor data.[52]

Ryans concluded that,

> Prediction of teacher effectiveness must be considered largely in the actuarial sense; individual prediction, as generally is the case in attempting to predict human behavior, is much more limited and is accomplished with a lesser degree of confidence.[53]

The most recent extensive summary of the teacher's personality and characteristics is that of Getzels and Jackson, who reviewed the literature from 1950 to 1963

---

[52] Ibid.

[53] Ibid.

and organized the investigations under the following headings:

1. Attitudes
2. Values, Interests, Favored Activities
3. Cognitive Abilities
4. Adjustment, Needs
5. Projective Techniques[54]

Attitudes

The most common instrument used to study teacher attitude is the Minnesota Teacher Attitude Inventory (M.T.A.I.). Generally, the results are interesting but not conclusive. While there seems to be no teacher attitude that is always highly correlated with teacher success, there were a number of findings which suggested some teacher characteristics:

1. The M.T.A.I. is able to predict how teachers will get along with their pupils.
2. The M.T.A.I. is able to differentiate between elementary academic field major and special field majors. It is also able to differentiate between elementary teachers educated in various types of institutions.
3. Guidance workers have higher M.T.A.I. scores than administrators and elementary teachers have higher scores than secondary teachers.
4. M.T.A.I. scores of teachers liked best by pupils are higher than scores of teachers liked least.

---

[54]Getzels and Jackson, op. cit., p. 507.

5. M.T.A.I. scores are higher after student teaching than before.

6. High M.T.A.I. scorers tend to be cooperative, friendly, objective, emotionally stable.

7. Low M.T.A.I. scorers seem to be subject to "fake good." Thus, some studies suggest that the use of the M.T.A.I. as a major factor in hiring a teacher or accepting a student for teacher training is not warranted in light of the instrument's susceptibility to faking.[55]

When Minnesota Multiphasic Personality Inventory (M.M.P.I.) scores are related to student teaching ratings, Moore and Cole[56] reported, "poorest" student teachers have higher mean T scores in each clinical scale. Flanagan[57] found that superior teachers rank higher in the MMPI scales of hypochondriasis, depression and masculinity. Lough found no significant statistical differences between two curriculum groups (elementary school teachers and music teachers) or between the total sample and the published norms and concluded,

> . . . The MMPI has little or no value in educational selection; it is not a useful instrument for differentiating between those who are more suited for one occupation than another.[58]

---

[55] Ibid., pp. 508-522.

[56] Clark H. Moore and David Cole, "The Relation of MMPI Scores to Practice Teaching Ratings," Journal of Educational Research, L (May, 1957). 711-716.

[57] Carroll E. Flanagan, "A Study of the Relationship of Scores of the MMPI to Success in Teaching as Indicated by Supervisory Ratings," Journal of Experiment Education, XXXIX (June, 1961), 329-354.

[58] Orpha M. Lough, "Women Students in Liberal Arts, Nursing, and Teacher Training Curricula and the MMPI," Journal of Applied Psychology, XXXI (August, 1947), 444.

Studies of other attitude measures such as the Authoritarianism (F) Scale indicate that teachers, as a group, are less authoritarian than other adults of similar status and that males are lower than female teachers on the F scale and classroom observations of authoritarianism.[59] Other measures have found that attitudes of high school teachers are less favorable toward children and toward current educational theory and practice than are the attitudes of teachers in the lower grades. Counselors seem to be better adjusted than teachers (or they are more aware of personality instruments and give more guarded responses). Elementary teachers have more favorable attitudes toward administrators, adult non-administrators, and pupils than do secondary teachers in both disguised and non-disguised inventories. Highly rated teachers were more favorable in their pupil and administrator attitudes and teaching behavior seemed to be related to teachers' attitudes towards pupils and administrators.[60]

Values, Interests, Favored Activities

Most studies of teachers' values, interests and favored activities used such instruments as the Allport-Vernon-Lindzey Study of Values, the Kuder Preference

---

[59] Getzels and Jackson, op. cit., pp. 522-523.

[60] Ibid., pp. 523-525.

Record and the Strong Vocational Interest Blank.[61] Such instruments have shown that:

1. Male education students are higher on Aesthetic and Social values and lower on Economic and Religious values than the norms.

2. Women are lower on Economic Values and higher in Theoretical and lower in Religion than the norms.

3. Women kindergarten teachers are higher in Aesthetic and Social values but lower in Economic values than are home economics teachers.

4. Male physical education teachers differ on all six value scores when compared to female physical education teachers.

5. Economic and Aesthetic values show a high correlation with teaching success.

6. Economic values show a negative correlation with student teaching success.

7. It seems that values indicative of student teaching success bear no relationship to values characteristic of the successful practicing teacher.

8. It appears that significant differences in values exist between teachers in different subject areas.[62]

Generally speaking measures of interest such as the Kuder and Strong do not seem to be suitable for predicting teaching success when all items are considered together. The Strong results have little or no relationship to teaching success. Successful student teachers seemed to be interested in working with people, selecting

---

[61] Ibid., pp. 525-533. [62] Ibid.

occupations involving teaching, and pursuing intellectual interests. "Poorest" student teachers avoided occupations related to teaching and selected those offering personal gain, considered salary most important, and were not interested in intellectual pursuits. However, these interests are not significantly correlated with either successful or non-successful teaching on the job.

Studies of the favored activities of teachers bear very little relationship to ratings of teacher effectiveness. Female teachers, however, tend to engage in school related activities, organizational activity and hobby activities. Males reported greater participation in sports. Single subjects reported greater activity in entertainment and subjects in higher teaching levels reported greater participation in out of school employment, sports and daily living activities. Lower teaching levels, on the other hand, reported greater participation in school related activities, organization and hobby activities. Subjects with higher salaries reported greater participation in the profession-related, school-related, daily living, organization and hobby activities.

In sum, there is very little we can say about the personality characteristics of the successful teacher on the basis of attitude, value, interest or favored activities studies. For example, as measured by the M.T.A.I., a teacher's school performance seems to be related to

his professional out of school activities; as measured by principal ratings, it seems to be related to his non-professional organization and sports out of school activities.

Cognitive Abilities

Studies of the cognitive abilities of teachers indicate that teachers as a group score appreciably higher on tests of intellectual ability than does the general population, but are somewhat below average when compared to most other professional groups. In most instances, the overlap between groups of teachers is too great to allow any definite statement concerning cognitive differences among various groups of teachers. Numerous tests of intelligence have been administered to teachers and related to one measure or another of teaching effectiveness. In general, the results are disappointing. In only 16 of more than 55 studies does the correlation coefficient exceed .30. Then again, 15 studies report negative correlations. Thus, it seems that no measure of cognitive ability is factorially related to marks obtained by subjects in practice teaching or on the job rating.[63]

---

[63] Ibid., pp. 570-574.

Adjustment and Needs

Many studies have been undertaken to investigate the adjustment and needs of teachers. Such instruments as the Minnesota Multiphasic Personality Inventory (MMPI), the Bell Adjustment Inventory, the Bernruter, the Washburne, the Thurstone and the Mooney Problem Check List have been used. Other than finding that the psychological health of teachers and student teachers is, on the whole, normal and average, the studies have not yielded any significant results for predicting teaching success.

Many studies conclude that the M.M.P.I. can detect no difference for education students and other students in other fields, nor do any of the standard scales differentiate between groups of teachers and non-teachers. M.M.P.I. tests administered before and after student teaching show no significant changes. In general, no standard score or combination of variables from the M.M.P.I. is able to successfully discriminate between "high" and "average" student teachers.[64]

Studies which utilized the Bell, the Bernruter, the Washburne, the Thurstone and the Mooney have not yielded any significant results for predicting success in teaching.[65]

---

[64] Ibid., pp. 534-545. [65] Ibid., pp. 545-547.

Although published studies using the Edwards Personality Preference Scale are as yet too few to justify any conclusions concerning the ultimate usefulness of the instrument in studying teacher personality, there have been some interesting findings:

1. Teacher education students score higher than liberal arts students on the following scales: deference, heterosexuality, order and endurance. Teacher education students scored low on exhibition.

2. Scores of the Edwards show that young female teachers more closely resemble their older, more experienced counterparts than did males.

3. The Edwards indicates that education students in a teachers college display personality configurations resembling those of practicing professionals far more than do education students in multipurpose institutions.[66]

A factor analytic study by Cook, Linden and McKay[67] using the Edwards Personal Preference Schedule and the Guilford-Zimmerman Temperament Survey identified several factors apparently related to effective teaching: docility, dependency authoritarianism, compulsive conformity and avoidance.

Studies using the Guilford-Zimmerman Temperament Survey with teachers, while not as numerous as those

---

[66] Getzels and Jackson, op. cit., pp. 545-546.

[67] Desmond L. Cook, et al., "A Factor Analysis of Teacher Trainee Responses to Selected Personality Inventories," Educational and Psychological Measurement, XXI (Winter, 1961), 865-872.

involving the M.M.P.I. do contribute to the total picture of teacher personality. Clark[68] found that teachers score higher than the general test norms and tend to be objective, agreeable, cooperative and appear in better mental health. Gowan and Gowan[69] reported that teacher candidates score higher than the general test norms in restraint, ascendancy, sociability, emotional stability, objectivity, friendliness and personal relations; Leeds[70] investigation confirmed the findings of Gowan and Gowan. Jones[71] offered a mildly affirmative answer to the question whether the Guilford test can be used to distinguish between "good" and "average" teachers.

Results with the Guilford test, although the small number of studies reduces the possibility of conflicting findings, add support to a psychologically favorable picture of the teacher.

---

[68]E. J. Clark, "The Mental Health of Elementary School Teachers as Measured by the Guilford-Martin Personality Battery," Paper read at the National Council on Measurements used in Education, Atlantic City, March, 1950.

[69]J. C. Gowan and May S. Gowan, "The Guilford-Zimmerman and the California Psychological Inventory in the Measurement of Teaching Candidates," California Journal of Educational Research, VI (January, 1955), 35-37.

[70]C. H. Leeds, "Teacher Attitude and Temperament as a Meaure of Teacher-Pupil Rapport," Journal of Applied Psychology, XXXX (October, 1956), 333-337.

[71]Margaret L. Jones, "Analysis of Certain Aspects of Teaching Ability," Journal of Experimental Education, XXV (December, 1956), 152-180.

> One is left, therefore, with the conclusion that, on the whole, teachers as a selected group are more likely to obtain "good" scores on the Guilford instrument than a random group. Reassuring as this may be, it does not seem to lead directly to a conceptual formulation of teacher personality.[72]

An instrument that has created research activity is Cattell's 16 P.F. Test, which has been applied to various groups of teachers. Lamke[73] found that teachers rated "good" and "poor" by administrative ratings differed on the 16 P.F. Test factors; "good" teachers

> are more than usually talkative, cheerful, placid, frank and quick; whereas poor teachers are below average in these respects . . . "good" teachers are above average in their tendencies to be gregarious, adventurous, frivolous, to have abundant emotional responses, strong artistic or sentimental interests, and to be interested in the opposite sex. The poor teachers are below average in these respects.[74]

Erickson[75] found that good teachers show practical concernedness, trustfulness, positive character and stability on the 16 P.F. factors, and Hadley's[76] study

---

[72] Getzels and Jackson, op. cit., p. 550.

[73] Lamke, op. cit., pp. 217-259.

[74] Ibid., p. 243.

[75] Harley E. Erickson, "A Factorial Study of Teaching Ability," Journal of Experimental Education, XXIII (September, 1954), 1-39.

[76] S. T. Hadley, "A Study of the Predictive Value of Several Variables to Student Teaching Success as Measured by Student Teaching Marks," Teacher's College Bulletin, LX, 3 (State Teachers College, Indiana, Pennsylvania, 1954), 1-10.

partially confirmed and partially contradicted Erickson's findings.

Other studies with the 16 P.F. Test report inconsistent and contradictory findings. Tarpey[77] reported that student teaching grades are related to personality factors measured by the 16 P.F. Test, whereas Montross,[78] who used two composite ratings of success in teaching found no significant correlation coefficients between 16 P.F. scores and the first measure of success. When related to the second measure of success, Montross reported only one of the 16 factor scores reached statistical significance.

> Too little has been done with the 16 P.F. Test, with teachers as subjects, to accept without question the results that have appeared and to make possible a comprehensive evaluation of the method. In some cases the findings of the several studies contradict one another.[79]

Thus, although it appears that much is known about the characteristics of good teachers, there is a dearth of clear cut results. The lack of empirical data to justify the use of any single personality measuring instrument suggests the need to take a different viewpoint on teacher characteristics.

---

[77]Sister M. S. Tarpey, "Personality Factors in Teacher Trainee Selection," British Journal of Educational Psychology, XXXV (June, 1965), 140-149.

[78]Harold W. Montross, "Temperament and Teaching Success," Journal of Experimental Education, XXIII (September, 1954), 34-97.

[79]Getzels and Jackson, op. cit., p. 553.

The need, for example, to take a perceptual view of teacher education has already been suggested by a number of writers.[80]

> Whether an individual will be an effective teacher depends upon the nature of his private world of perceptions. It follows that the perceptual world of the student must be a matter of vital concern to teacher education programs.
>
> If teacher education is to be concerned with changing student perceptions, we need clear definitions of what the perceptual organization of effective teachers are like.[81]

Projective Techniques

One of the most promising approaches in appraising the perceptual organization of teacher candidates is the use of projective techniques which includes instruments as the Rorschach, Thematic Apperception Test, Holtzman Inkblots, Word Association Tests and Incomplete Sentences.

---

[80] W. R. Dixon and W. C. Morse, "The Prediction of Teaching Performance: Empathic Potential," Journal of Teacher Education, XII (September, 1961), 322-329; Margaret Lindsey, New Horizons for the Teaching Profession (Washington, D.C.: National Commission on Teacher Education and Professional Standards, National Education Association, 1961); W. W. Lynch, "Person Perception: Its Role in Teaching," Indiana University School of Education Bulletin, XXXVII (1962), 1-37; P. E. McClendon, "Teacher Perception and Working Climate," Educational Leadership, XX (November, 1962), 104-108; C. P. Ramsey, "Leadership Perception," Educational Leadership, XX (December, 1962), 151-154; Ryans, op. cit.

[81] Combs, The Professional Education of Teachers, p. 19.

Getzels and Jackson[82] reviewed the projective technique literature and reported the major findings of projective instrument studies with teachers:

The Rorschach

1. There is a positive correlation of .61 between adjustment scores on the Rorschach and teaching success.[83]
2. Emotionally constricted persons are most likely to be found in a group of teachers that pupils like least.[84]
3. There is a positive relationship between absence of neurotic signs and favorable pupil ratings.[85]
4. A comparison of Rorschach performance and student teaching evaluation shows that among elementary student teachers, a pattern of three Rorschach scores differentiated between "desirable" and "undesirable" student teachers. These scores, W, M, and C, reflect a strong drive, desire for achievement, emotional outgoingness with interest in the environment, people, liability and suggestability.[86]

---

[82] Getzels and Jackson, op. cit., p. 554.

[83] Ibid.

[84] Ibid., p. 555.

[85] Ibid.

[86] Ibid., p. 556.

5. Among secondary student teachers, no Rorschach patterns were consistently associated with any supervisory evaluation.[87]

In a quantitative investigation of the Rorschach Inkblot Test as applied to student teachers, Sultan[88] found that the test appears to elicit some variables of general psychological significance, but cannot be accepted as an all-round technique for assessing personality in student teachers.

Lewis[89] conducted a study using the Group Form of the Rorschach, The Structured-Objective Rorschach Test (SORT), with 283 student teachers and attempted to identify those clusters of personality attributes measured by the SORT that contributed to success in student teaching. Lewis concluded that the student teachers studied were better adjusted than the normative population, and that of the 25 SORT variables only two factors, Practical and Deduction, were positively correlated with success in teaching.

---

[87] Ibid.

[88] E. E. Sultan, "A Quantitative Investigation of the Rorschach Inkblot Test As Applied to Student Teachers," British Journal of Social and Clinical Psychology, IV (September, 1965), 197-206.

[89] James N. Lewis, "The Relationship of Attributes Measured by the Structured-Objective Rorschach Test and Success in Student Teaching" (unpublished Ph.D. dissertation, North Texas State University, 1966).

Thematic Apperception Test

Although the Thematic Apperception Test (TAT) itself has apparently not been used in research on teaching, several modifications of the TAT technique are reported in the literature.

The Adult-Child Interaction Test (ACI), Alexander[90] reported, may be used to predict certain types of teacher behavior which agree closely with observed classroom behavior. Ohlsen and Schultz also used the ACI in an attempt to discriminate between good and poor student teachers and found that the instrument could not significantly discriminate between the groups. Ohlsen and Schultz reported however:

> . . . the fantasy production of . . . good student teachers contained more "school oriented" elements than did those of the poorer group. This phenomenon although admittedly somewhat vague, serves to introduce a theme that gains in insistence as the findings from studies using other projective techniques are examined.[91]

"All we can say is that this approach merits further study."[92]

---

[90]T. Alexander, "The Adult-Child Interaction Test: A Projective Test for Use in Research," Monograph Social Research in Child Development, XVII, No. 2 (1952) (Whole No. 55).

[91]Ohlsen and Schultz, op. cit., pp. 18-27.

[92]Getzels and Jackson, op. cit., p. 558.

Word Association and Sentence Completion Tests

The theme that Ohlsen and Schultz referred to is reported in a study by Goodenough, Fuller and Olson[93] who used a list of 238 stimulus words to which subjects responded by writing the first word that came to them after reading the stimulus word. An analysis of the results which focused only on the responses was able to discriminate between education students and liberal arts students besides significantly distinguishing between student teachers who receive high ratings from those who did not. Perhaps the most noteworthy findings were that successful student teachers replied to the stimulus words with more educationally oriented words than did poorer student teachers. The education group made references to children's activities at least five times as frequently as liberal arts students and successful student teachers responded to neatness, opposites, and coordinates at least twice as often as liberal arts students.

Hilton[94] used a modification of the classic word association method in a study of 130 teacher trainees. His instrument, the Word Completion Form, contains 108 items, each consisting of a series of connected letters from which several common words could be made by adding

---

[93]Goodenough, Fuller and Olson, op. cit.

[94]Hilton, op. cit.

letters (e.g., GOA _____; CHI _____:). The score on the instrument was the number of "education" words used as completions. Hilton found that successful student teachers produced far more "educationally oriented" words than did least successful student teachers.

> Although the relationships were not strong, they pointed to the potential fruitfulness of searching for "educational" content in the unrestricted cognitive products of teachers.[95]

The findings of the studies by Goodenough, Fuller and Olson, Hilton, and Ohlsen and Schultz bear an interesting relationship to each other. The results suggest that the unrestricted cognitive products of teachers contain more "educational elements" than do comparable products from non-teacher groups. Further, the amount of "educational elements" seems also to distinguish those rated successful in teaching from those rated unsuccessful.

> The phenomenon common to these studies . . . suggests that teachers characteristically view their world--provide it with structure and meaning--in ways that distinguish them from non-teaching groups. Further, this perceptual mode appears related to measures of the teacher's effectiveness. The fact that students in training show this perceptual tendency suggests that it is not merely a product of teaching experience . . . the role that such a perceptual orientation might play in the decision to become a teacher seems worthy of further research.[96]

---

[95] Getzels and Jackson, op. cit., p. 561.

[96] Ibid., p. 566.

As a technique, research with projective instruments appears to be the most practical way to investigate perceptual orientation, and the most widely recognized projective technique is the Rorschach. However, several Rorschach studies reported in the literature suggested that the test suffers from incurable limitations. Cronbach,[97] Zubin,[98] and Hertz[99] wrote that major problems in the Rorschach arise from examiner-subject interaction and variations in style of inquiry, lack of satisfactory internal consistency or test-retest reliability for many scores, lack of agreement in scoring criteria for several variables, and from widely varying number of responses often obtained for the ten Rorschach cards. The complex, curvilinear relation between number of responses and most other scores on the Rorschach make it impossible to establish adequate norms for most Rorschach scores. "The Rorschach displays an alarming variety of psychometric deficiencies."[100] Thus it appears

---

[97]L. J. Cronbach, "Statistical Methods Applied to Rorschach Scores: A Review," Psychological Bulletin, XXXXVI (September, 1949), 393-429.

[98]J. Zubin, "Failure of the Rorschach Technique," Journal of Projective Techniques, XVIII (September, 1954), 303-315.

[99]Marguerite R. Hertz, "The Use and Misuse of the Rorschach Method," Journal of Projective Techniques, XXIII (March, 1959), 33-48.

[100]Richard W. Cohan, "The Holtzman Inkblot Test," in The Sixth Mental Measurements Yearbook, ed. by O. K. Buros (3rd ed.; Highland Park, New Jersey: The Gryphon Press, 1965), p. 439.

that although the Rorschach has a rich and extensive history in projective investigations, the findings are always subject to some criticism.

Research of the Holtzman Technique

The Holtzman Inkblot Instrument

However, within recent years a projective instrument has appeared which purports to retain the advantages of the Rorschach while minimizing its statistical limitations.

> Drawing heavily upon studies with the Rorschach, the Holtzman Inkblot Technique (HIT) is a new projective method designed to overcome psychometric limitations in the Rorschach by constructing completely new sets of inkblots. Unlike the Rorschach which has only ten inkblots in a single form, the HIT consists of two parallel forms, A and B, each of which contains forty-five inkblots constituting the test series and two practice blots, X and Y which are identical in both forms. Thus, standardized responses can be obtained from a total of 92 different inkblots rather than just ten.[101]

The HIT also differs from the Rorschach in other respects than merely the number of inkblots:

1. The characteristics of the HIT stimuli are richer and more varied in color, form and shading.

---

[101]Wayne H. Holtzman, "The Holtzman Inkblot Technique," in Projective Techniques in Personality Assessment, ed. by Albert I. Rabin (New York: Springer Publishing Company, 1968), p. 136.

2. The blots vary considerably in degree of symmetry or balance, providing a new stimulus dimension for analysis.

3. The subject is encouraged to give only one response per card rather than as many as he wishes, thereby holding constant the number of responses given.

4. A brief, simple inquiry follows immediately after each response.

5. Carefully matched, parallel forms of the HIT are available, permitting the use of test-retest designs and the study of change within the individual.

6. Standardized percentile norms are provided for 22 inkblot scores on a variety of populations, facilitating interpretation and analysis.

7. Group methods of administration and computer scoring make it possible to use the HIT for rapid, large-scale screening as well as individual diagnosis and assessment.[102]

The original standardization studies for the individual HIT involved fifteen different samples and over 1500 protocols. Percentile norms were established for eight reference groups including college students, average adults, seventh graders, elementary school children, five year olds, chronic schizophrenics, mental retardates, and depressed patients.

> Several reliability studies for the different groups are reported. In the main, the obtained coefficients are acceptably high and some are remarkably so, although on some of the variables, they are too low to warrant confidence. Validity studies have dealt primarily with group differences and relationships with other techniques. In the developmental, cognitive,

---

[102] *Ibid.*

> and perceptual aspects of personality functioning, concurrent validity seems reasonably satisfactory. For the identification of psychopathology, the variables of rejection, form appropriateness, movement, pathognomic verbalization, integration, human and popular are especially powerful.[103]

Since publication of the HIT monograph[104] and its accompanying materials in 1961, a considerable amount of information has accumulated on reliability, validity, basic experimental investigations, and other exploratory work. Many investigations have been conducted to determine relationships between HIT scores and independently obtained behavioral, personality, cognitive, perceptual, developmental, sociocultural and psychodiagnostic measures.

Correlations between HIT scores and standard measures of intelligence, scholastic achievement, and convergent thinking are low though statistically significant.[105] Tests of divergent thinking, creativity, and other forms of cognitive functioning also correlate with inkblot scores.[106]

Although various paper and pencil approaches to personality study by self-inventory have been studied

---

[103]William H. Thetford, "The Holtzman Inkblot Technique," in Sixth Mental Measurements Yearbook, ed. by O. K. Buros (3rd ed.; Highland Park, New Jersey: The Gryphon Press, 1965), p. 443.

[104]Wayne Holtzman, et al., Inkblot Perception and Personality (Austin: University of Texas Press, 1961), p. 417.

[105]Ibid., p. 154. [106]Ibid.

for HIT correlates, the results are without much success. And similar poor results were obtained for the MMPI and the Edwards Personal Preference Schedule when related to HIT variables.[107] One major exception, however, is the study reported by Richter and Winter[108] who used the intuitive-perceptive scales of the Myers-Briggs Type indicator for defining high and low groups of creative women. They found that women with a high amount of this type of creativity had significantly higher scores on nine HIT variables.

A number of studies, although limited, report the effectiveness of the HIT in classifying groups and differential diagnosis, that is, studies in which behavioral ratings of manifestations play a determining role in the criterion classification. Although sociometric peer ratings within normal groups generally fail to correlate significantly with HIT scores, ratings by skilled observers do show significant correlations with the inkblot variables.[109]

Several major studies have been reported relating HIT scores and developmental trends in children. Highly

---

[107] Ibid., p. 155.

[108] H. Richter and W. D. Winter, "Holtzman Inkblot Correlates of Creative Potential," Journal of Projective Techniques and Personality Assessment, XXX (February, 1966), 62-67.

[109] Ibid., p. 156.

significant age trends for all but two of the individual scores and for many of the pattern scores were found across four of the normal populations included in the HIT standardization studies.[110]

Other HIT investigations show that the instrument is particularly useful in differential diagnosis in psychopathic disorders and because of the non-verbal nature of the inkblots, the HIT has been used in major cross-cultural studies ranging from industrial societies to primitive, non-literate tribes.[111]

The Group HIT

In a series of studies which employed large numbers of college students a standardized group method of administration was developed[112] and shown to yield inkblot scores highly comparable to the standard, individual version of the HIT. Of the twenty-two inkblot variables in the standard HIT, only reaction time is clearly lost in the group method. Space and Balance were dropped because they rarely occur. Form Appropriateness, and Pathognomic Verbalization were omitted because they present specific scoring problems in the group administration method.

---

[110] Ibid., p. 157.

[111] Ibid., pp. 158-162.

[112] J. D. Swartz and W. H. Holtzman, "Group Method of Administration for the Holtzman Inkblot Technique," Journal of Clinical Psychology, XIX (October, 1963), 433-441.

The group version consists of showing slides of the HIT to subjects and asking them to record what they see on an answer sheet. Thus, the group method of administration makes it possible to collect a large number of test protocols from a wide range of populations in a very economical and efficient manner.

Perhaps the most important question which could be asked concerning a group administered test is to what extent does the method produce scores which are equivalent to individually administered HIT records? In order to answer this question, the individual HIT was administered to 418 college students under standard test conditions described in the test manual and, shortly thereafter, the group HIT was administered to the same sample. Table 6 presents a comparison of the mean scores and standard deviations on the individual and group HIT. Only five of the fifteen variables showed a significant shift in mean scores across the two methods. Location and Color yielded higher mean scores for the group method than for the individual, while the reverse proved true for Barrier and Popular.

> The increase in scores for Location and Color . . . can be attributed to . . . different instructions used by the examiner. In the group method, the examiner used trial blots X and Y to illustrate responses employing detailed areas and color. In the individual method such

explicit instruction is not necessary since the examiner writes down what the subject says.[113]

TABLE 6.--Means and standard deviations for 15 variables from individual inkblot technique and a similar set of 15 variables from group inkblot technique.[114]

INDIVIDUALLY ADMINISTERED HIT			GROUP ADMINISTERED HIT	
Mean	S.D.	VARIABLE	Mean	S.D.
29.8	13.7	L	38.3	16.1
82.1	15.1	R	80.0	12.7
19.1	9.8	FD	26.3	11.5
10.2	5.7	C	11.5	7.2
37.1	14.3	Sh	32.2	13.3
6.5	3.8	M	6.0	3.7
22.7	8.9	I	25.8	8.7
24.0	6.7	H	23.4	6.3
10.8	4.8	A	12.1	7.7
10.6	5.4	At	11.1	5.8
7.9	3.7	Ax	6.2	3.4
4.1	2.9	Hs	4.0	2.8
10.4	3.1	Br	8.8	2.7
1.2	2.3	Pn	1.5	2.7
3.3	2.7	P	3.5	2.8

The shift in mean scores for Popular was probably due to subtle changes in the stimulus when transformed from inkblots on a card to inkblot slides and the slight drop in Barrier was probably due to the decreased elaboration by subjects who write out their response. Means

[113] Ibid., p. 445.

[114] W. H. Holtzman, et al. "Comparison of the Group Method and the Standard Individual Version of the Holtzman Inkblot Technique," Journal of Clinical Psychology, XIX (October, 1963), 441.

for the remaining variables were remarkably similar across the two methods.

Thus,

> the high degree of comparability of the group method and the standard individual method suggests that the group version may be confidentially substituted for the more time-consuming individual version.[115]

The group method, however, represents no saving of time in scoring since the process is a laborious one requiring about thirty minutes for each protocol. This led to a demand for an economical, time-saving and objective scoring system.

The group of HIT researchers pioneered the development of a computer scoring system for group administered HIT protocols. These studies, and the computer scoring system, reported in Chapter I, pages 19-24, suggest that computer scored group HIT records are satisfactorily equivalent to hand scoring methods.

As a final test to determine the extent to which computer scored group administered HIT records are equivalent to hand scored individually administered records, Gorham replicated a former study by Holtzman.[116] In the original study, 100 college students took the group HIT one week after taking the individual HIT. The individually administered hand-scored records were available and the protocols for the group records were

---

[115] Ibid., p. 448.

[116] Ibid., pp. 441-449.

obtained. Both sets of records were keypunched and computer scored. When computer scoring was substituted for hand-scoring for the group administered records, the correlations were essentially those obtained in the original study and are presented in Table 7.

During the past three years a large number of populations have been sampled and extensive norms have been developed using the group method of administration and computer scoring. Gorham[117] has recently published the computer scored norms for over 5000 subjects in three general classes: normal subjects, clinical subjects and cultural samples.

Within the normal subject classification, Gorham reported six major samples: 7th graders, 9th graders, 11th graders, college sophomores, Navy enlistees and Michigan State University student teachers. The present study provided the norms for the largest single sample in the computer scored norms: 802 Michigan State student teachers.

Since the time consuming and technical difficulties of administering and scoring individual Holtzman protocols is impractical in teacher education selection programs, the group administered, computer scored method was used

---

[117]Gorham, et al., "Norms for the Computer Scored Holtzman Inkblot Technique," pp. 1279-1305.

in this study of perceptual modes (i.e. Holtzman variables) related to success in student teaching, sex, and major teaching field.

TABLE 7.--Correlation coefficients between hand scored individual and group HIT records and computer scored individual and group HIT records.

Hand Scored*[118]	Variable	Computer Scored[119]
r		r
.69	L	.67
.57	R	.57
.60	FD	.56
.42	C	.33
.31	Sh	.38
.56	M	.55
.23	I	.22
.46	H	.43
.53	A	.54
.36	At	.32
.44	Ax	.42
.47	Hs	.37
.34	Br	.42
.27	Pn	.24
.14	P	.28

*Computer scored reliability coefficient found by Holtzman et al. These values are not corrected for attenuation due to test-retest unreliability since the comparison of computer and hand scoring is sufficiently demonstrated by the uncorrected values for both methods.

[118] Ibid., p. 444.

[119] Gorham, "Validity and Reliability Studies of a Computer Based . . . ," pp. 65-70.

CHAPTER III

PRESENTATION AND DISCUSSION OF THE DATA

All of the hypotheses of the study were statistically investigated using the computer facilities at Michigan State University and the resources of the College of Education Research Bureau. The decision to reject the null hypotheses was arbitrarily set at the five per cent level of significance; that is, if there were differences among each of the mean Holtzman Inkblot variable scores of student teachers grouped according to rated success, sex, and teaching field, at the five per cent level of significance, the null hypotheses was rejected.

The computer scored Holtzman variable scores for each student teacher were obtained from the Holtzman Inkblot Test Scoring Service at Silver Springs, Maryland, and analyzed for the overall means and standard deviations as presented in Table 8.

Table 9 presents selected Holtzman variable scores and their corresponding percentile ranks for the sample and represents norms for Michigan State student teachers.

TABLE 8.--Overall mean scores and standard deviations of the Holtzman Inkblot variable scores for the student teacher population (N = 802).

Variable	Mean	Standard Deviation
L	31.57	13.14
R	1.00	2.11
FD	82.46	12.02
C	10.43	5.78
Sh	4.25	2.53
M	29.63	11.58
I	5.01	1.94
H	27.37	6.58
A	28.28	6.02
At	2.59	2.46
Sx	0.70	1.06
Ab	1.45	1.58
Ax	8.23	4.18
Hs	8.81	4.21
Br	10.00	3.48
Pn	3.38	2.12
P	8.61	2.50

TABLE 9.--Selected Holtzman variable scores and corresponding percentiles of Michigan State University student teachers.

Variables	Percentiles										
	99	90	80	70	60	50	40	30	20	10	1
Selected Holtzman Variable Scores (N = 802)											
L	62	48	42	38	35	32	29	25	20	14	12
R	11	3	2	1	0	0	0	0	0	0	0
FD	106	97	92	89	86	83	80	76	72	66	47
C	28	19	15	13	11	10	8	7	5	4	1
Sh	11	7	6	5	4	4	3	3	2	1	0
M	59	45	39	35	31	29	26	23	20	15	8
I	10	8	7	6	5	5	4	4	3	3	1
H	43	36	33	31	29	27	26	24	22	19	12
A	43	36	33	31	30	28	27	25	23	20	14
At	10	5	4	3	3	2	2	1	1	0	0
Sx	5	2	1	1	1	0	0	0	0	0	0
Ab	8	4	3	2	1	1	1	0	0	0	0
Ax	20	14	11	10	9	8	7	6	5	3	1
Hs	20	14	12	11	9	8	7	6	5	4	1
Br	19	14	13	12	11	10	9	8	7	5	2
Pn	9	6	5	4	4	3	3	2	1	1	0
P	15	12	11	10	9	9	8	7	7	5	3

Presentation of the Data

Hypothesis One

The first hypothesis was that there will be no significant differences among each of the mean Holtzman variable scores for student teachers grouped according to rated success in student teaching. Table 11 presents the mean scores, standard deviations and F statistic and its approximate significance, of the student teacher population classified according to rated success.

An examination of Table 10 shows that there were no statistically significant differences among the mean scores on sixteen of the seventeen Holtzman variables. The only variable that did reach statistical significance at the designated five per cent level was R (rejection). However, this finding is somewhat questionable since inspection of the data for the R variance revealed that the assumption of approximately equal variances or homoscedasticity on which the F test is based, was not entirely justified.

Such responses as "just looks like an inkblot," "nothing," or "must be something there but I can't see it" as well as not responding at all were counted as rejections. Responses of "a bunch of colors" or "smear of paint" were borderline responses that were not considered as rejections.

Table 10.--Means, standard deviations, F statistic and its approximate significance of Holtzman Inkblot variables among student teacher groups classified according to rated success.

Rating:	OUTSTANDING N = 51		GOOD N = 232		AVERAGE N = 432		BELOW AVG. N = 74		POOR N = 13		F	Approx.
Variable	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.	Statistic	Significance
L	31.80	14.01	30.93	14.17	31.98	12.92	31.36	10.50	29.92	12.60	0.304	0.875
R	1.92	4.02	0.68	1.51	1.00	2.00	1.12	1.91	2.07	3.54	4.887	0.001*
FD	80.58	9.93	82.24	12.41	82.53	12.27	83.48	9.31	85.53	17.50	0.667	0.608
C	10.17	6.26	10.31	5.41	10.53	5.96	10.17	5.86	11.84	4.20	0.307	0.873
Sh	4.39	2.89	4.44	2.53	4.18	2.55	3.95	2.14	4.07	2.39	0.719	0.579
M	27.47	11.16	29.75	11.52	29.96	11.99	28.98	9.55	28.46	11.63	0.631	0.640
I	4.72	1.99	4.90	1.95	5.06	1.91	5.20	1.92	5.61	2.56	1.041	0.385
H	26.05	7.15	27.36	6.05	27.32	6.66	28.21	6.79	29.76	8.97	1.249	0.288
A	27.17	6.59	28.25	5.76	28.35	6.10	28.85	6.03	27.38	5.76	0.683	0.604
At	2.74	2.27	2.59	2.90	2.56	2.17	2.55	2.60	2.92	3.20	0.121	0.975
Sx	0.64	0.79	0.66	1.08	0.73	1.04	0.71	1.17	0.69	1.70	0.199	0.939
Ab	1.25	1.23	1.64	1.73	1.39	1.56	1.41	1.46	0.84	0.80	1.727	0.142
Ax	8.31	3.93	8.34	4.04	8.34	4.34	7.10	3.44	8.46	5.39	1.277	0.277
Hs	8.68	3.43	8.56	3.75	9.09	4.53	7.95	4.04	9.30	4.28	1.480	0.206
Br	9.62	3.61	10.33	3.46	9.89	3.40	9.79	3.94	10.38	2.84	0.879	0.475
Pn	3.68	2.15	3.39	2.13	3.39	2.11	3.05	2.08	3.69	2.39	0.774	0.542
P	8.35	2.55	8.48	2.46	8.71	2.51	8.68	2.50	8.46	2.75	0.510	0.728

*Significant at the 5% level.

Since the R score was the only probable significant finding, the null hypothesis was rejected for the R variable but may not be rejected for the remaining sixteen variables.

Hypothesis Two

The second hypothesis was that there will be no significant differences among each of the mean Holtzman variable scores for the male student teacher group and the female student teacher group. Table 11 presents the mean scores, standard deviations, and F statistic and its approximate significance, of student teachers classified according to sex.

An examination of Table 11 shows that the mean Holtzman scores on nine variables significantly differed between the sexes. Males and females scored significantly different on the R (rejection), FD (form definiteness), M (movement), I (integration), A (animal), At (anatomy), H (human), Sx (sex) and P (popular) variables.

As previously indicated, the R (rejection) score was the total number of inkblots to which the subject did not give a scorable response. From the data in Table 11, it is evident that males gave a significantly higher number of responses classified as rejections than did females, subject, of course, to the assumption of equal variances.

TABLE 11.--Means, standard deviations, F statistic and its approximate significance of Holtzman Inkblot variable scores among student teachers grouped according to sex.

Variable	MALES N = 171 Mean	MALES N = 171 S.D.	FEMALES N = 631 Mean	FEMALES N = 631 S.D.	F Statistic	Approx. Significance
L	32.05	13.01	31.45	13.18	0.281	0.603
R	1.71	2.54	0.81	1.93	25.310	0.0005*
FD	78.46	13.49	83.54	11.36	24.758	0.0005*
C	9.73	5.50	10.62	5.80	3.228	0.069
Sh	4.38	2.67	4.21	2.49	0.546	0.466
M	25.78	10.53	30.67	11.64	24.627	0.0005*
I	4.36	1.91	5.19	1.91	25.010	0.0005*
H	26.02	6.85	27.74	6.46	9.298	0.003*
A	26.30	6.11	28.82	5.89	24.130	0.0005*
At	3.22	2.95	2.41	2.28	14.531	0.0005*
Sx	1.01	1.30	0.62	0.94	17.997	0.0005*
Ab	1.49	1.66	1.43	1.56	0.146	0.702
Ax	8.38	4.16	8.19	4.19	0.263	0.614
Hs	8.73	4.07	8.83	4.25	0.082	0.767
Br	9.59	3.69	10.11	3.41	2.942	0.083
Pn	3.56	2.26	3.33	2.08	1.593	0.204
P	7.42	2.48	8.94	2.40	52.614	0.0005*

*Significant at the 5% level.

The sexes also differed significantly on FD (form definiteness) scores which were responses that were definite in form, regardless of the goodness of fit to the inkblot. For example, responses such as "splattered paint" or "a cloud" were concepts that have an indefinite or non-specific form and received a lower Form Definiteness score. A highly structured concept such as "a boy riding a bicycle" received a higher FD score. Form Definiteness was scored on a five point scale, ranging from zero (concepts that were completely formless or lacking in specificity) to four (concepts that were highly definite in form).

From the data presented in Table 11 it is evident that females, as a group, scored significantly higher than males on the Form Definiteness variable.

Females also scored significantly different from males on the M (movement) score. The Movement variable reflects the degree of movement, tension or dynamic energy projected into the inkblot. The scoring system for Movement is a five point scale as follows:

0 - no movement, or static potential for movement.

1 - static potential for movement as indicated by such participles as sitting, looking, resting, lying.

2 - casual movement, such as walking, talking, climbing, reaching.

3 - dynamic movement such as lifting, dancing, running, weeping.

4 - violent movement such as whirling, exploding.[1]

Responses such as "two men reclining" and "a dog gazing into the sky" received Movement scores of 1; responses such as "whirling dervishes" or "an atomic bomb blast" received Movement scores of 4.

Since females scored significantly higher (mean of 30.67) than males (mean of 25.78) one may conclude that females characteristically projected more responses that received Movement scores than males.

Besides scoring differently on R, FD and M, the sexes also differed significantly on the I (integration) scores. Holtzman wrote that,

> The organization of two or more adequately perceived blot elements into a larger whole is necessary for a score of 1 on Integration. The particular elements that are integrated may involve either identical or non-identical contents and the resulting concept need not utilize the whole inkblot.[2]

Although Phillips, Kaden and Waldman[3] distinguished among several types of Integration (i.e., functional, collective, positional, structural) credit for Integration

---

[1]Holtzman, et al., Inkblot Perception and Personality, p. 51.

[2]Ibid., pp. 64-65

[3]L. Phillips, S. Kaden, and M. Waldman, "Rorschach Indices of Developmental Level," Journal of Genetic Psychology, LXXXXIV (1959), 267-285.

was either given (0) or not given (1), thus permitting a theoretical range of 0 to 45 across the forty-five inkblots. Responses such as "roosters fighting," "a collection of shells," "a candle stuck in the neck of a bottle" and "a bee standing up" received a score of 1. Females had a significantly higher mean score (5.19) than males (4.36) so it is evident that females characteristically gave more integrated responses than males.

Table 11 also shows that the sexes differed significantly in the content scores of H (human), A (animal), At (anatomy), and Sx (sex). For each of these content areas a three point weighting system permitted the scaling of content as follows:

0 - No reference to humans, animals, penetration, or sex.

1 - References to parts of humans, animals, X rays, medical drawings and socially accepted sexual activity and expressions (thighs, bust, pregnant woman, kissing, embracing, etc.).

2 - References to whole humans, animals, visceral and crude anatomy responses, blatant sex (penis, vagina, anus, breast, etc.).[4]

From an analysis of the content scores presented in Table 11, it is evident that females gave significantly more responses with H (human) and A (animal) content than did males. On the other hand, males scored significantly

---

[4]Holtzman, et al., Inkblot Perception and Personality, p. 68.

higher than females in At (anatomy) and Sx (sex) content.

Males also scored significantly different from females on the P (popular) variable. A Popular response is one which occurred frequently enough in the original HIT normative samples to be classified as popular among normal subjects. For a response to be judged Popular, it must have occurred in at least one out of seven protocols in the original normative protocols. This arbitrary cutting point was chosen, Holtzman reported,[5] because it provided a sufficient number of inkblots with Populars to yield a variable with good psychometric qualities. Several inkblots so closely resemble human or animal form, for example, that most people perceive these concepts. In other words, the inkblot has a "pull" for a certain content response; responding like most people do yields a score of 1 on the Popular variable. In the group Holtzman Test, there are twenty-five inkblots scorable for Popular resulting in a theoretical Popular range of 0 to 25.

Table 11 shows that females characteristically perceived significantly more Populars than males.

In sum, males and females differed significantly on nine of the seventeen Holtzman variable scores: R (rejection), FD (form definiteness), M (movement), I (integration),

---

[5] Ibid., p. 80.

A (animal), At (anatomy), H (human), Sx (sex) and P (popular). Thus, the null hypothesis was rejected for these variables but was not rejected for the remaining eight variables of L (location), C (color), Sh (shading), Ab (abstract), Ax (anxiety), Hs (hostility), Br (barrier) and Pn (penetration).

Since there were significant variable score differences between the sexes, it was logical for the investigator to undertake a re-examination of the first hypothesis to determine if initial grouping by sex and then an analysis of mean scores related to success would alter the failure to reject the first null hypothesis.

In other words, Hypothesis One stated:

> There will be no significant differences among each of the mean Holtzman variable scores of student teachers classified according to rated success,

and was modified to:

> (1,A) There will be no significant differences among each of the mean Holtzman variable scores of MALE student teachers classified according to rated success, and,
>
> (1,B) There will be no significant differences among each of the mean Holtzman variable scores of FEMALE student teachers classified according to rated success.

Table 12 presents the means, standard deviations, and F statistic and its approximate significance, of male student teachers grouped according to rated success. An examination of Table 12 shows that there are no statistically significant differences among sixteen of seventeen

TABLE 12.--Means, standard deviations, F statistic and its approximate significance of Holtzman Inkblot variable scores among male student teacher groups classified according to rated success.

Rating:	OUTSTANDING N = 10		GOOD N = 43		AVERAGE N = 99		BELOW AVG. N = 19		POOR N = 0		F	Approx.
Variable	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.	Statistic	Significance
L	33.20	12.38	28.69	15.04	33.36	12.78	32.21	8.17	--	--	1.322	0.269
R	4.20	4.77	1.09	1.74	1.68	2.47	1.94	2.27	--	--	4.315	0.006*
FD	78.80	10.30	77.32	12.96	78.50	14.70	80.63	9.50	--	--	0.264	0.891
C	6.50	4.32	9.65	4.82	10.37	6.13	8.26	3.98	--	--	2.028	0.112
Sh	5.20	3.61	4.53	2.89	4.32	2.56	3.89	2.25	--	--	0.578	0.630
M	25.30	11.78	23.86	9.96	26.54	10.91	26.47	9.28	--	--	0.680	0.565
I	3.70	2.16	4.00	1.71	4.55	1.97	4.57	1.80	--	--	1.339	0.263
H	24.78	9.41	25.27	6.20	26.20	7.01	27.47	6.08	--	--	0.594	0.619
A	24.10	6.53	27.06	4.81	26.12	6.61	26.68	5.88	--	--	0.707	0.549
At	3.10	2.23	3.44	4.20	3.29	2.49	2.42	1.98	--	--	0.565	0.639
Sx	0.80	1.03	1.13	1.83	1.01	1.20	0.84	1.38	--	--	0.287	0.834
Ab	0.80	0.63	1.86	2.31	1.46	1.43	1.15	1.21	--	--	1.565	0.200
Ax	8.10	3.78	9.44	4.41	8.03	4.18	7.94	3.48	--	--	1.252	0.292
Hs	7.20	2.57	8.41	3.60	9.00	4.32	8.84	4.46	--	--	0.698	0.554
Br	9.00	3.05	10.02	3.94	9.53	3.67	9.26	3.67	--	--	0.334	0.800
Pn	3.40	2.67	3.48	2.35	3.62	2.26	3.52	2.09	--	--	0.058	0.981
P	7.20	3.58	7.20	2.17	7.55	2.49	7.36	2.58	--	--	0.226	0.878

*Significant at the 5% level.

mean scores on the Holtzman variables. R (rejection) was the only variable that did reach statistical significance at the five per cent level. But this was the only significant finding assuming homoscedasticity and was not supported by significant differences among the other variables. Thus, null hypothesis 1,A may not be rejected for the remaining sixteen variables, and the conclusion is that mean Holtzman variable scores did not differ significantly among male student teachers grouped according to rated success for sixteen of the seventeen Holtzman variables.

Table 13 presents the means, standard deviations, and F statistic and its approximate significance, of female student teachers grouped according to rated success. An examination of the table shows that there were no significant differences for sixteen of the seventeen variables. However, like the males, R (rejection) was the only variable that did reach significance if one assumes homoscedasticity. But since this finding was not supported by other statistically significant findings, null hypothesis 1,B may not be rejected for sixteen of the seventeen variables.

Hypothesis Three

The third hypothesis stated that there will be no significant differences among each of the mean Holtzman

TABLE 13.--Means, standard deviations, F statistic and its approximate significance of Holtzman Inkblot variable scores among female student teacher groups classified according to rated success.

Rating:	OUTSTANDING N = 41		GOOD N = 189		AVERAGE N = 333		BELOW AVG. N = 55		POOR N = 13		F	Approx.
Variable	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.	Statistic	Significance
L	31.46	14.50	31.43	13.95	31.57	12.95	31.07	11.25	29.92	12.60	0.062	0.993
R	1.36	3.67	0.58	1.44	0.79	1.79	0.96	1.73	2.07	3.54	2.998	0.018*
FD	81.02	9.92	83.36	12.03	83.72	11.20	84.47	9.13	85.53	17.50	0.727	0.573
C	11.07	6.37	10.47	5.53	10.57	5.91	10.83	6.27	11.84	4.20	0.258	0.904
Sh	4.19	2.70	4.42	2.44	4.14	2.55	3.98	2.12	4.07	2.39	0.538	0.708
M	28.00	11.09	31.09	11.45	30.98	12.12	29.85	9.57	28.46	11.63	0.847	0.496
I	4.97	1.89	5.10	1.95	5.21	1.87	5.41	1.93	5.61	2.56	0.590	0.670
H	26.39	6.59	27.83	5.93	27.66	6.52	28.47	7.06	29.76	8.97	0.967	0.425
A	27.92	6.47	28.52	5.93	29.02	5.79	29.60	5.95	27.38	5.76	0.882	0.474
At	2.65	2.30	2.39	2.48	2.35	2.01	2.60	2.79	2.92	3.20	0.427	0.789
Sx	0.60	0.73	0.56	0.79	0.65	0.97	0.67	1.10	0.69	1.70	0.354	0.841
Ab	1.36	1.31	1.59	1.57	1.36	1.60	1.50	1.53	0.84	0.80	1.170	0.320
Ax	8.36	4.01	8.08	3.92	8.43	4.39	6.92	3.42	8.46	5.39	1.591	0.175
Hs	9.04	3.54	8.60	3.80	9.11	4.60	7.65	3.88	9.30	4.28	1.638	0.163
Br	9.78	3.75	10.40	3.36	9.99	3.32	9.98	4.04	10.38	2.84	0.570	0.684
Pn	3.75	2.04	3.37	2.08	3.32	2.07	2.89	2.07	3.69	2.39	1.162	0.326
P	8.63	2.21	8.77	2.43	9.06	2.41	9.14	2.33	8.46	2.75	0.839	0.501

*Significant at the 5% level.

Inkblot variable scores of student teachers grouped according to major teaching fields or grade levels.

To adequately test this hypothesis, the thirty-six subject/grade classifications presented in Table 14 were collapsed into thirteen major subject/grade areas since in several of the thirty-six teaching fields there were few subjects. Table 13 illustrates the combined sample size for the major teaching areas. For example, student teachers in French (N=18), German (N=1), Spanish (N=12), Russian (N=4) and Latin (N=2) were grouped under the major teaching area of language (N=37); student teachers in biology (N=25), chemistry and physics (N=5), beginning and advanced mathematics (N=38), and general science (N=5) were grouped in the major teaching area of science and mathematics (N=72). Junior high school student teachers were not included in the test of this hypothesis since student teachers at this grade level may be in any teaching area and as such would tend to confuse the data.

Table 15 presents the means, standard deviations, and F statistic and its approximate significance, of the Holtzman variable scores of student teachers grouped according to major teaching field or grade level. An examination of Table 15 shows that the mean Holtzman variable scores on eleven variables differed significantly among the major teaching areas. That is, student teachers grouped according to major teaching field scored

TABLE 14.--The student teacher population (excluding junior high school student teachers) classified according to major teaching field and sex.

Specific Teaching Field	Number of Student Teachers			Major Teaching Field	Resultant Sample Size		
	Male	Female	Total		Male	Female	Total
Elementary							
Grades: K,1,2,	1	112	113	Elementary Education	5	230	235
Grades: 3,4	1	70	71				
Grades: 5,6	3	48	51				
Physical Educ.	20	17	37	Physical Educ.	20	17	37
French	1	17	18	Languages	6	31	37
German	0	1	1				
Spanish	2	10	12				
Russian	2	2	4				
Latin	1	1	2				
English	12	59	71	English	15	72	87
Speech	13	13	16				
Biology	15	10	25	Science and Math	40	32	72
Chem./Physics	4	1	5				
Beginning Math	6	10	16				
Advanced Math	13	8	21				
General Science	2	3	5				
Distrib. Educ.	4	1	5	Business Education	4	11	15
Business (Sec'y)	0	10	10				
Home Economics	0	45	45	Home Economics	0	45	45
Industrial Arts	13	0	13	Industrial Arts	13	0	13
Instrumental Mus.	7	2	9	Music Education	8	3	11
Vocal Music	1	1	2				
Agriculture	15	0	15	Agriculture	15	0	15
Soc. Studies/ Eng. Combined	3	7	10	History/ Social Studies	28	61	89
History/Soc.St.	25	52	77				
Geography	0	2	2				
Art	6	30	36	Art	6	30	36
Special Educ.				Special Education	5	64	69
Sph. Correct'n							
Elementary	1	23	24				
Secondary	0	7	7				
Visually Handicapped	0	4	4				
Homebound	1	7	8				
Deaf	0	2	2				
Emotionally Dis.	0	4	4				
Mentally H'cap	3	17	20				
TOTALS	165	596	761	TOTALS	165	596	761

TABLE 15.--Means, standard deviations, F statistic and its approximate significance of Holtzman variable scores of student teachers classified according to major teaching field.

Major Teaching Field		Elem. Educ. N=[illegible]	Phys. Educ. N=[illegible]	Lang-uages N=[illegible]	Eng-lish N=[illegible]	Science & Math N=[illegible]	Busi-ness N=[illegible]	Home Econ. N=[illegible]	Industrial Arts N=13
Variable									
L	Mean	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	59.84
	S.D.	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	15.52
F	Mean	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	1.84
	S.D.	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	2.67
Fr	Mean	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	67.69
	S.D.	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	11.67
K	Mean	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	11.15
	S.D.	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	5.69
Sc	Mean	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	6.07
	S.D.	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	3.81
Mf	Mean	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	21.76
	S.D.	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	13.00
Pt	Mean	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	4.07
	S.D.	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	2.46
Hy	Mean	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	25.07
	S.D.	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	9.10
Pa	Mean	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	24.07
	S.D.	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	3.61
At	Mean	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	3.30
	S.D.	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	2.42
Sx	Mean	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	0.92
	S.D.	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	0.64
Ab	Mean	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	1.23
	S.D.	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	1.92
Ax	Mean	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	6.61
	S.D.	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	5.65
Hs	Mean	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	8.00
	S.D.	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	5.52
Hy	Mean	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	8.07
	S.D.	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	4.40
Pd	Mean	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	3.76
	S.D.	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	3.00
P	Mean	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	6.53
	S.D.	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	2.29

*Significant at the 5% level.

Music Educ. N=11	Agri-culture N=15	Social Sciences N=89	Art N=36	Special Educ. N=69	F Statistic	Approximate Significance
30.45 12.21	40.06 10.22	32.98 12.52	22.47 13.89	27.31 12.05	3.843	0.0005*
1.36 2.01	1.66 2.12	1.84 3.49	0.77 1.47	0.81 2.12	3.101	0.0005*
81.54 12.72	77.93 10.93	32.50 13.60	83.75 12.51	85.00 10.75	3.080	0.0005*
8.28 4.66	9.80 6.51	9.56 5.94	11.88 6.75	10.84 5.69	1.015	0.432
3.63 2.37	4.60 2.55	3.82 2.28	3.27 2.07	4.63 2.53	2.086	0.016*
27.63 8.78	19.40 7.35	28.89 12.04	35.72 13.43	36.13 11.10	6.160	0.0005*
4.27 1.42	3.46 1.80	5.39 2.04	5.61 2.32	5.65 1.58	3.344	0.0005*
26.09 5.41	21.80 7.08	28.55 6.42	30.02 5.43	27.60 6.04	2.451	0.004*
28.00 5.63	29.06 7.07	26.82 5.80	30.44 5.17	28.72 6.13	4.544	0.0005*
2.45 2.54	2.80 2.90	2.59 2.20	2.55 3.13	2.00 1.68	1.649	0.074
0.27 0.46	0.40 0.63	0.61 0.84	1.08 1.74	0.68 0.86	2.629	0.002*
2.27 2.24	0.93 0.79	1.53 1.63	1.66 1.60	1.65 1.78	1.304	0.211
8.63 4.78	7.13 3.33	7.83 3.75	8.41 3.94	9.07 4.51	1.398	0.161
8.72 3.52	8.00 3.52	9.11 4.11	9.33 4.57	10.04 4.88	1.753	0.052
10.54 3.36	9.20 3.62	10.20 3.78	9.50 3.90	10.60 3.29	2.504	0.003*
2.63 1.68	3.06 1.98	3.05 2.04	3.72 2.26	3.60 1.93	0.741	0.711
6.27 2.10	7.06 2.31	8.39 2.33	9.66 2.26	9.05 2.44	3.565	0.0005*

significantly different on the L (location), R (rejection), FD (form definiteness), Sh (shading), M (movement), I (integration), H (human), A (animal), Sx (sex), Br (barrier), and P (popular) variables. Several of these variables (R, FD, M, I, H, A, Sx and P) were also significantly different between the male and female student teacher groups and were discussed on pages 88 through 97. The remaining variable scores (L, Sh and Br) although not significantly different between the sexes, were significantly different among student teachers grouped according to major teaching field.

The L (location) variable score represented the tendency to break down the inkblot perception into smaller fragments. For example, a subject received an L score of zero each time all or most of the inkblot was used in the perception. If, on the other hand, the perception was a large area (an entire side, for example) but something less than the entire blot, the L score for that perception was one. If an area less than one-half of the inkblot was used, the L score was two. Thus, the total L score was computed by summing the individual L scores across the forty-five inkblots yielding a theoretical L range of 0 to 90.

The Sh (shading) Holtzman variable score represented the apparent primacy or importance of shading in the perception and was rated on a three point scale as follows:

0 - shading not used as a determinant

1 - shading used only in a secondary manner as an elaboration of the percept reported (similar to Fc, FK, Fy in the Rorschach)

2 - Shading used as a primary determinant with no form present (similar to cF, c, K, YF, Y in the Rorschach)

A response such as "mist or fog" uses shading as a primary determinant and, as such, received a Sh score of two. Thus, the theoretical scoring range of the Sh variable is 0 to 90.

The concept of the Holtzman Br (barrier) variable referred to any protective covering, membrane, shell or skin that might be symbolically related to the perception of body-image boundaries. A score of one was given to each response where Barrier was present, a score of zero was given when Barrier was absent. Responses such as "woman in a high-necked dress" or references to things that are armored or much dependent on their own containing walls for protection such as a "tank" or "man in armor" received a Barrier score of one. References to things being covered, surrounded or concealed also received a score of one.

Since student teachers grouped according to major teaching field scored significantly different on L, R, FD, Sh, M, I, H, A, Sx, Br, and P, the conclusion is that null hypothesis III was rejected for these variables, but was not rejected for the remaining variables

of C, At, Ab, Ax, Hs and P. Since, however, there were significant variable score differences between the sexes (hypothesis II, pages 96-97), the investigator undertook an examination of hypothesis III to determine if initial grouping by sex and then classification into major teaching fields would alter the findings for hypothesis III. In other words, hypothesis III stated:

> there will be no significant differences among each of the mean Holtzman Inkblot variable scores of student teachers grouped according to major teaching fields or grade levels

and was modified to:

> (III,A) There will be no significant differences among each of the mean Holtzman Inkblot variable scores of MALE student teachers grouped according to major teaching field or grade levels, and,
>
> (III,B) There will be no significant differences among each of the mean Holtzman Inkblot variable scores of FEMALE student teachers grouped according to major teaching fields or grade levels.

Hypothesis III-A

Table 16 presents the means, standard deviations, and F statistic and its approximate significance, of male student teachers grouped according to major teaching field or grade level. An examination of Table 16 shows that there are no statistically significant differences among sixteen of the seventeen mean variable scores. At (anatomy) was the only variable that did reach statistical significance at the five per cent level.

TABLE 16.--Means, standard deviations, F statistic and its approximate significance of Holtzman variable scores of male student teachers classified according to major teaching field.

Major Teaching Field		Elem. Educ. N=5	Phys. Educ. N=20	Lang-uages N=6	Eng-lish N=15	Science & Math N=40	Busi-ness N=4	Home Econ. N=0	Industrial Arts N=13
Variable									
L	Mean	33.20	37.70	21.16	32.46	29.55	28.75		29.84
	S.D.	17.71	11.97	16.85	13.53	11.84	16.25		15.52
R	Mean	1.60	2.20	2.00	0.80	1.75	3.00		1.84
	S.D.	1.94	3.47	3.09	1.14	2.33	3.55		2.67
FD	Mean	82.80	74.45	90.16	81.20	81.27	78.25		67.69
	S.D.	13.12	11.00	10.68	12.21	14.14	16.45		11.67
C	Mean	8.20	9.45	8.00	10.33	9.45	12.25		11.15
	S.D.	2.77	5.39	7.37	5.08	6.13	4.99		5.69
Sh	Mean	4.20	3.80	4.66	4.73	3.90	5.00		6.07
	S.D.	3.34	2.48	2.58	2.60	2.52	1.41		3.81
M	Mean	23.40	23.70	31.16	28.66	26.67	24.25		21.76
	S.D.	10.66	11.03	8.90	10.95	10.11	11.02		13.00
I	Mean	3.80	3.60	5.16	4.60	4.55	3.75		4.07
	S.D.	1.48	1.72	1.16	1.80	2.02	1.50		2.46
H	Mean	24.20	23.95	27.00	27.56	25.97	24.00		25.07
	S.D.	4.54	8.04	5.17	5.65	7.15	5.29		9.10
A	Mean	28.00	25.70	27.66	24.33	26.40	26.50		24.07
	S.D.	5.83	4.85	5.50	6.66	7.07	8.58		3.61
At	Mean	7.80	4.15	4.83	2.53	3.30	1.75		3.30
	S.D.	10.75	2.51	3.06	1.84	2.51	1.25		2.42
Sx	Mean	2.40	1.20	0.83	1.00	1.55	1.00		0.92
	S.D.	4.27	1.28	0.75	1.30	1.76	1.41		0.64
Ab	Mean	1.60	1.25	1.83	1.80	1.37	2.50		1.23
	S.D.	1.14	1.55	2.22	1.82	1.23	4.35		1.92
Ax	Mean	9.60	7.45	8.33	10.06	8.92	9.00		8.61
	S.D.	5.81	4.79	5.71	4.68	3.91	3.36		5.65
Hs	Mean	6.40	7.75	8.50	10.40	9.20	10.50		8.00
	S.D.	2.70	3.76	5.16	3.13	4.65	3.41		5.52
Br	Mean	7.60	9.95	10.83	11.13	9.32	9.75		8.07
	S.D.	1.67	3.18	4.62	2.29	3.89	1.70		4.40
Pn	Mean	2.40	3.00	4.16	4.06	4.07	5.25		3.76
	S.D.	1.34	2.55	4.07	1.79	2.41	2.21		3.00
P	Mean	9.00	6.95	8.83	7.20	7.24	7.50		6.53
	S.D.	2.73	2.35	1.94	1.89	2.86	1.29		2.29

*Significant at the 5% level.

Music Educ. N=8	Agri-culture N=15	Social Sciences N=28	Art N=6	Special Educ. N=5	F Statistic	Approximate Significance
28.00	40.06	32.78	31.50	23.20	1.843	0.051
13.58	10.22	10.34	16.37	6.26		
1.75	1.66	1.85	1.83	1.20	0.344	0.974
2.25	2.12	3.34	1.83	0.83		
78.62	77.93	78.78	79.50	83.40	1.777	0.062
13.35	10.93	13.26	17.73	12.17		
8.12	9.80	9.00	13.66	8.80	0.636	0.796
5.11	6.51	5.42	4.84	3.56		
4.00	4.60	4.14	3.33	4.80	0.865	0.576
2.61	2.55	2.39	2.06	3.56		
29.12	19.40	25.75	31.66	33.00	1.522	0.128
9.90	7.35	9.30	15.29	11.76		
4.50	3.46	5.10	4.83	4.20	1.301	0.229
1.19	1.80	1.74	2.99	1.09		
27.62	21.80	29.17	28.00	25.80	1.557	0.117
3.58	7.08	5.87	6.78	4.43		
25.87	29.06	26.17	29.66	26.80	0.813	0.627
4.96	7.07	5.91	4.63	7.52		
2.50	2.80	2.96	1.50	1.60	2.107	0.023*
2.82	2.90	1.81	1.22	1.14		
0.37	0.40	0.60	0.50	0.80	1.806	0.057
0.51	0.63	0.56	0.54	1.30		
2.50	0.93	1.71	2.00	1.20	0.794	0.645
2.56	0.79	1.78	1.26	0.83		
8.62	7.13	7.53	7.16	9.80	0.746	0.692
4.47	3.33	3.14	3.12	2.04		
9.50	8.00	8.39	8.50	10.40	0.808	0.632
3.66	3.52	3.34	4.84	2.40		
10.37	9.20	9.89	9.00	11.00	0.780	0.659
3.88	3.62	4.62	3.52	1.58		
3.12	3.06	3.00	3.33	3.40	0.968	0.477
1.64	1.98	1.98	1.63	0.54		
6.50	7.06	8.21	8.16	7.20	1.015	0.436
1.85	2.31	2.72	2.31	2.86		

The At variable is a content variable previously discussed since this variable score was also significantly different between the sexes (page 90). But this was the only significant finding and was not supported by significant differences among the other variables. Thus, null hypothesis III,A may not be rejected for the remaining sixteen variables and the conclusion is that mean Holtzman variable scores did not differ significantly among male student teachers grouped according to major teaching fields for sixteen of the seventeen variables.

Hypothesis III-B

Table 17 presents the means, standard deviations, and F statistic and its approximate significance, of female student teachers grouped according to major teaching field or grade level. An examination of Table 17 shows that there are ten variables, L (location), R (rejection), Sh (shading), M (movement), I (integration), A (animal), Sx (sex), Hs (hostility), Br (barrier), and P (popular) that are significantly different among the female student teachers grouped according to major teaching fields or grade levels. Each of these variables has been discussed previously in other hypotheses (L, location, page 104; R, rejection, page 88; Sh, shading, page 104; M, movement, page 92; I, integration, page 93; A, animal, page 94; Sx, sex, page 96; and Br, barrier, page 105) except for the Hs or Hostility variable.

TABLE 17.--Means, standard deviations, F statistic and its approximate significance of Holtzman variable scores of female student teachers classified according to major teaching field.

Major Teaching Field		Elem. Educ. N=23	Phys. Educ. N=17	Lang-uages N=31	Eng-lish N=72	Science & Math N=32	Busi-ness N=11	Home Econ. N=45	Industrial Arts N=0
Variable									
L	Mean	32.72	35.00	30.12	30.48	36.12	25.27	33.33	
	S.D.	12.60	11.01	15.02	12.87	13.10	14.05	14.55	
R	Mean	0.65	0.47	1.16	0.75	1.21	0.54	0.28	
	S.D.	1.57	0.62	1.71	1.68	2.43	1.03	0.69	
FD	Mean	83.60	83.41	82.22	84.27	84.15	76.09	82.64	
	S.D.	10.07	12.95	14.22	10.07	13.80	9.71	13.87	
C	Mean	10.72	9.58	11.51	9.47	11.03	11.72	11.37	
	S.D.	5.56	5.63	7.65	5.22	6.55	7.21	5.00	
Sh	Mean	4.28	4.58	4.32	3.83	4.00	4.36	4.97	
	S.D.	2.37	1.69	1.92	2.33	2.43	2.90	2.97	
M	Mean	30.63	28.52	23.51	30.22	26.06	32.00	32.24	
	S.D.	11.31	11.77	9.23	10.15	13.20	13.60	10.09	
I	Mean	4.93	4.94	4.83	5.22	5.13	5.36	5.51	
	S.D.	1.80	2.16	1.88	1.88	2.03	2.15	2.20	
H	Mean	26.92	29.64	26.90	28.47	27.25	27.90	28.80	
	S.D.	6.21	7.49	6.64	6.72	7.50	6.04	6.73	
A	Mean	30.17	27.47	27.03	27.26	28.56	26.45	28.60	
	S.D.	5.56	6.18	5.29	5.80	6.83	4.13	6.54	
At	Mean	2.47	3.17	3.06	2.27	1.87	2.00	2.22	
	S.D.	2.04	2.21	3.95	1.77	1.40	2.72	2.38	
Sx	Mean	0.50	0.82	0.58	0.66	0.56	0.36	0.55	
	S.D.	0.70	1.00	0.67	0.76	0.80	0.67	0.96	
Ab	Mean	1.30	1.88	1.29	1.50	0.96	1.54	1.93	
	S.D.	1.50	1.79	1.34	1.54	1.09	1.21	1.88	
Ax	Mean	8.19	6.88	7.45	9.30	6.34	7.81	7.80	
	S.D.	4.12	3.37	4.74	4.03	3.98	4.75	3.62	
Hs	Mean	8.65	6.41	7.61	9.44	7.53	7.18	9.04	
	S.D.	3.95	2.98	3.79	4.14	4.63	3.68	4.56	
Br	Mean	9.59	10.70	10.83	11.00	8.84	11.36	10.75	
	S.D.	3.38	3.51	3.63	2.96	2.86	4.03	3.83	
Pn	Mean	3.29	3.17	3.41	3.55	2.65	2.90	3.28	
	S.D.	2.04	2.24	1.96	2.03	1.94	1.97	2.30	
P	Mean	8.80	9.41	8.00	9.29	9.28	9.09	8.71	
	S.D.	2.38	2.31	2.11	2.48	2.55	2.77	2.56	

*Significant at the 5% level.

Music Educ. N=3	Agri-culture N=0	Social Sciences N=61	Art N=30	Special Educ. N=64	F Statistic	Approximate Significance
37.00		33.08	20.66	27.64	3.937	0.0005*
3.46		13.49	12.90	12.36		
0.33		1.83	0.56	0.78	2.564	0.005*
0.57		3.59	1.33	2.19		
89.33		84.21	84.60	85.12	0.804	0.625
7.76		13.52	11.41	10.73		
8.33		9.81	11.53	11.00	0.794	0.634
4.16		6.19	7.08	5.81		
2.66		3.67	3.26	4.62	1.849	0.050*
1.52		2.24	2.11	2.47		
23.66		30.34	36.53	36.37	4.407	0.0005*
3.21		12.92	13.16	11.11		
3.66		5.52	5.76	5.76	1.877	0.046*
2.08		2.17	2.19	1.57		
22.00		28.26	30.43	27.75	1.590	0.106
8.18		6.76	5.16	6.15		
33.66		27.11	30.60	28.87	3.363	0.0005*
2.51		5.91	5.34	6.05		
2.33		2.42	2.76	2.03	1.005	0.437
2.08		2.35	3.37	1.72		
0.00		0.62	1.20	0.67	2.078	0.024*
0.00		0.95	1.88	0.83		
1.66		1.45	1.60	1.68	1.259	0.250
1.15		1.57	1.67	1.83		
8.66		7.96	8.66	9.01	1.774	0.062
6.65		4.02	4.08	4.65		
6.66		9.44	9.50	10.01	2.264	0.013*
2.51		4.41	4.58	5.03		
11.00		10.34	9.60	10.57	2.196	0.017*
0.00		3.36	4.03	3.39		
1.33		3.08	3.80	3.62	1.137	0.331
1.15		2.09	2.38	2.00		
5.66		8.47	9.96	9.20	2.324	0.011*
3.05		2.14	2.17	2.37		

The Hostility variable score was based on a four point scale which considered symbolic, implicit or explicit signs of hostility in the response. The more explicit the hostility, the higher the Hs score, with a range of zero for no hostility to three for direct hostility among humans or animals seen in violent destructive action toward each other with elaboration of gore, injuries, blood, death, etc.

Since female student teachers grouped according to major teaching field scored significantly different on L, R, Sh, M, I, A, Sx, Hs, Br, and P, null hypothesis III,B was rejected for these variables but was not rejected for the remaining variables of FD, C, H, At, Ab, Ax, and Pn.

Summary of the Findings

Table 18 presents a summary of the finding for each variable as it differentiates among student teachers grouped according to success (hypotheses I; I,A; I,B), sex (hypothesis II) and major teaching field (hypotheses III; III,A; III,B).

With the entire student teacher sample grouped according to success (hypothesis I), only R (rejection) appeared to significantly differentiate among the success ratings. However, the unequal variance of the R variable (or the failure of the R variance to meet the assumption

TABLE 18.--Summary of findings.

	Hypotheses	Variables																
		L	R	FD	C	Sh	M	I	H	A	At	Sx	Ab	Ax	Hs	Br	Pn	P
I	Student Teachers Grouped According to Rated Success	NS	*	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS
I,A	Male Student Teachers Grouped According to Rated Success	NS	*	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS
I,B	Female Student Teachers Grouped According to Rated Success	NS	*	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS
II	Student Teachers Grouped According to Sex	NS	*	*	NS	NS	*	*	*	*	*	*	NS	NS	NS	NS	NS	*
III	Student Teachers Grouped According to Major Teaching Field	*	*	*	NS	*	*	*	*	*	NS	*	NS	NS	NS	*	NS	*
III,A	Male Student Teachers Grouped According to Major Teaching Field	NS	NS	NS	NS	NS	NS	NS	NS	NS	*	NS	NS	NS	NS	NS	NS	NS
III,B	Female Student Teachers Grouped According to Major Teaching Field	*	*	NS	NS	*	*	*	NS	*	NS	*	NS	NS	*	*	NS	*

NS = Not significant at the five per cent level or less
* = Significant at the five per cent level or less

of homoscedasticity for the F test) suggested that the significance of the R variable is doubtful. When the student teachers were first grouped according to sex (hypotheses I,A; I,B) and then according to success rating, the R variable was the only significant finding, although (and similar to the hypothesis I finding for R), the unequal variance suggested a questionable significant finding.

When the student teachers were grouped according to sex (hypothesis II) and the variables studied for significant differences between males and females, nine variables (R, FD, M, I, H, A, At, Sx, and P) were found to be significantly different.

Finally, when the sample was classified according to major teaching field (hypothesis III), eleven variables (L, R, FD, Sh, M, I, H, A, Sx, Br, and P) were found to be significantly different among teaching fields; when the student teachers were first classified according to sex and then teaching field (hypotheses III,A and III,B) only one variable, At, significantly differentiated among males classified according to teaching fields. For females grouped according to teaching fields, however, ten variables (L, R, Sh, M, I, A, Sx, Hs, Br, and P) were found to be significantly different.

Discussion of the Findings

In the immediately preceding section of the present study, the major hypotheses were statistically analyzed and conclusions drawn. It has been the major focus of the study to determine what relationship, if any, exists between Holtzman Inkblot variable scores and success in student teaching, sex, and major teaching field. The hypotheses were necessarily limited in that, prior to this study, no relationships were known to exist. And in the investigation several significant findings were reported. To predict, hypothesize or guess what the significant differences mean, however, or to say what they are due to is, at best, speculative.

The fundamental principle of clinical interpretation consists of constant formulation of hypotheses, and as more evidence for a given hypothesis accumulates, more confidence is placed in it. As evidence appears to the contrary, the hypothesis is modified, changed or otherwise eliminated entirely from consideration. Evidence may come from clinical tests, case histories, diagnostic interviews and a wealth of other psychological techniques. The types of statements made in this section are based on the whole backlog of dynamic personality theory and psychoanalytic research and rely heavily on published Rorschach technique interpretative materials. The statements are not presented as facts, but simply as basic or first-level

hypotheses. If one looks upon the following discussion in this light, the findings are seen simply as hypotheses derived from an essentially group or non-clinical setting and should be considered as only gross or basic statements of assessment of the results. The writer feels compelled to make the point repeatedly of the dangers of interpreting findings without other supporting clinical evidence. Further verification of the comments are necessary before one can minimize the tenuousness of any given statement. In other words, Rabin's[4] admonition of the "shoestring operation" (daring and detailed interpretation without sufficient information) should be pointed out to those who would infer more than can be substantiated and is apropos of the ensuing discussion.

In addition, even if individual diagnostic interviews and case histories were available for comparative purposes, it should be noted that many hypotheses derived from projective techniques are not capable of being verified or discredited by knowledge of overt behavior. It has been pointed out by many clinicians that it is merely interesting if hypotheses from a projective technique and a case history agree, but highly important when they do not. The implication is that projective

---

[4]Albert I. Rabin, Varieties of Rorschach Interpretation, paper delivered in symposium on "The Uses and Misuses of the Rorschach," at the Michigan Psychological Association meetings in Detroit, May, 1958.

techniques tap levels of personality not revealed in the ordinary interview situation and as such make a real contribution to personality assessment.

Hypothesis One--Discussion of Significant Findings

Hypothesis one stated that there would be no significant relationships among each of the mean Holtzman variable scores of student teachers grouped according to rated success. For sixteen of seventeen variables, no significant differences existed. However, one variable, R (rejection), did significantly differentiate rated success in student teaching among the student teachers. However the significance of the R variable in this and the other hypotheses is questionable since the statistical test of significance among mean scores, the F test, assumes equal variances or homoscedasticity. Cursory investigation of the R variance (Table 10) shows a variance range from 1.51 to 4.02 thus suggesting that the equal variance assumption for the F test is not justified in this instance. Nevertheless, a discussion of the R variable does seem appropriate.

There are many methods of rejection ranging from no response at all (answer sheet left blank) to responses such as "just an inkblot" or "nothing at all." It is especially hazardous to interpret the meaning of a high rejection score on a group test. It could be that the

subject lost his place on the answer sheet; it could be that the subject was momentarily distracted, or, perhaps, that his pen simply ran out of ink while the test was in progress. If, however, one assumes that the subject saw the inkblot and failed to see anything at all, several tentative hypotheses, subject to further verification, are possible.

High rejection scores are typically made by subjects who are relatively non-responsive to their environments and/or are psychologically defensive. One possible interpretation, then, is that student teachers who score at the extreme ends of the rating scale are less defensive and more responsive to their environments than others. However, this single significant finding suggests that such an interpretation is very doubtful at best since it is not supported by additional significant differences in other variable scores. The mean R scores of outstanding and poor student teachers are approximately the same, thus, the possible significance of the R variable loses its meaningfulness in differentiating between outstanding and poor student teachers.

Hypothesis Two--Discussion of Significant Findings

Hypothesis two investigated the mean Holtzman scores of student teachers grouped according to sex and confirmed

that male and female variable scores differed significantly on nine variables: R, FD, M, I, H, A, At, Sex, and P.

The R (rejection) variable, besides being related to success in student teaching, is also significantly related to the sex of the subject, although there is some question of whether the assumption of homoscedasticity was met. Males gave responses scored as Rejections twice as often as females and, considering the previous discussion on the Rejection variable, this suggests that males, as a group, may be less responsive or more guarded in their perceptions than females. On the other hand, it may simply be that the male student teachers were uncooperative.

On the FD (form definiteness) variable, females appeared to perceive concepts with definite structure more often than males, suggesting that females are more demanding than males for definite concepts and less given to perceiving abstract concepts that do not have structure. Such a notion is doubtful at best, however, since a subject who is critical and demanding with regard to acceptable definiteness may well reject an inkblot, and males have a significantly higher rejection score than females.

On the other hand, if one assumes that the higher male rejection scores were not due to a demand for form

definiteness but to momentary distractions or other causes, one may hypothesize that, as a group, females are less given to perceiving abstract, free-form concepts than males. Such is certainly not a conclusion, but only a hypothesis based on an assumption.

As a group, females also scored higher on the M (movement) variable than males. Psychologically, movement projection is typically related to creativity, fantasy, self control, imagination, good intelligence and perhaps the capacity for delay of gratification. Again, an interpretation is not warranted, but one may safely conclude that, as a group, males project less energy into the inkblots than females, although the exact meaning of a high or low M score lies in individual analysis and supporting clinical evidence.

Besides scoring significantly higher in FD and M from males, females also scored higher on the I (integration) variable. Integration scores represent the organization of two or more perceived blot elements into a larger whole. For example, "two roosters fighting" or "a collection of shells" represent integrations or organized perceptions which are somewhat different from instantaneous perceptions. Psychologically, the subject sees the whole blot, analyzes it into components, sees a relationship between or among the parts and organizes the inkblot into a whole. Thus, the process is one of

higher analytic and organizational sophistication suggesting that female student teachers tended to synthesize related blot stimuli into organized perceptions more often than males.

The remaining variables (H, human; A, animal; At, anatomy; Sx, sex; and P, popular) are primarily content oriented and are related to the response content. For example, "two people talking" has human content; "a large elephant" is animal content, etc.

Assessing the significant differences in response content between any groups is extremely tentative, particularly with results from a group test. The data are only meaningful in an actuarial sense and do not lend themselves to easy interpretation. If, for example, many fierce human figures are perceived by a particular group, what does this mean? Are females acting out their aggression? Are they inhibiting it, repressing it? Perhaps they are having a reaction formation against it. It is ludicrous to suggest that females, as a group, act out their aggression.

In addition, the meaning of seeing many humans, for example, is unclear unless one knows the type of human or what they are doing. "Two humans talking" has a different psychological meaning than "two humans kneeling." In a sense, then, one can only deduce content meaning from an individual analysis of a protocol, and group

differences in content can only be commented upon in the most general sense.

Thus, it appears that there is no easy method to assess the meaning of the content differences other than through individual comparisons of projection with conscious attitudes such as may be obtained from case histories or from responses to other projective techniques. And even then the meaning may elude the clinician. What follows are comments about what one can say generally about the differences in content scores, although the meanings of the group differences lie beyond the scope of this study.

In the Holtzman Inkblots, females, as a group, perceived significantly more human figures than males. Generally, one could say then that the female student teachers seem to be more "people oriented" than males, since perception of human figures is related to identification. On an individual basis, however, if a subject is afraid of people, he may not see many human figures. Pre-occupation with particular parts of the body may point to certain concerns or difficulties, or he may not be able to determine whether human figures he sees in the blots are males or females, revealing problems of sexual identification. But these comments are individual in nature, based on individual protocol analysis, and nothing of this nature can be said about the group as a whole

other than females appear to be more strongly identified with other people than males.

Much of the same can be said for Animal content. Females perceived more animal figures than did males, but its meaning lies in individual analysis. Animals are perceived by subjects of all ages, both normal and abnormal. These responses may represent potential adjustive techniques. Even if the scores were to indicate what kind of animal, their meaning is still largely speculative, although there is some suggestion that seeing fierce animals such as tigers and lions, may refer to certain aggressive tendencies in the individual that the subject is trying to handle in some way. Passive animals, such as cows and sheep, and domestic animals such as dogs or cats may show passivity and dependent attitudes. Since the data show that females perceived more animals than males, but do not indicate what kind of animal, we can only cite the fact but cannot specify its meaning.

The same interpretive difficulties occur in assessing Sexual and Anatomy responses. Males scored significantly higher in both of these areas than did females. The meaning attached to sex responses needs to be studied in relation to other factors in the individual's test protocol and with some knowledge of the subject's life history. It is impossible to say on the basis of sexual responses alone whether the individual is avoiding sexual

problems by naming them, or whether the subject is worried about sexual problems. If it is impossible to deduce the meaning on an individual basis without other confirming data, so much the more to comment on group data. Generally, though, one could say that the males are more "sexually oriented" and appeared to have more of an interest in sexual themes than females.

Anatomy responses, too, are exceedingly difficult to interpret especially from a group test where individual case histories are not available. Typically, Anatomy responses are often given by subjects who are trying to show off intellectually as one way of covering up real feelings. Anatomy responses may also indicate feelings of intellectual inadequacy, especially when the responses avoid any commitment to definite form or specification. Thus, it could be speculated that males, as a group, are less prone to show their real feelings than females.

The P (popular) variable, however, is somewhat easier to interpret than other content areas. The P variable is not comparable to other content area scores since the Popular variable indicates how often the subjects see things as other people do. On this basis, it can be said that female student teachers, when compared to male student teachers, see things more often as other people do. This may be related to sociological needs and pressure to conform. In other words, the Popular score reveals

the group's tendency to view the world in the same way as most other people. Since females see significantly more popular perceptions than males, one can characterize the group as having a need to think as other people do, or at least more than male student teachers do.

Hypotheses I,A and I,B

Since the investigation of hypothesis II identified several variable scores which differentiated significantly between the sexes, and since hypothesis I did not identify any variable scores (other than R) which were significantly different among student teachers grouped according to success ratings, the investigator re-examined hypothesis I by grouping the student teachers by sex before seeking significant differences when classified according to success ratings.

Hypothesis I,A stated that there would be no significant differences among each of the mean Holtzman variable scores of MALE student teachers grouped according to success rating; hypothesis I,B stated the same but for FEMALE student teachers. For sixteen of the seventeen variables in both hypotheses no significant differences were found. The single exception, however, was the R variable, which was discussed previously on page 117 and need not be re-discussed here.

Hypothesis III

Hypothesis III investigated the mean Holtzman scores of student teachers grouped according to major teaching field and identified eleven variables (L, location; R, rejection; FD, form definiteness; Sh, shading; M, movement; I, integration; H, human; A, animal; Sx, sex; Br, barrier; and P, popular) which were significantly different among the thirteen major teaching areas. Eight of these variables, R, FD, M, I, H, A, Sx and P were also significantly different between the sexes and were discussed on pages 117 through 125.

L, Sh, and Br, however, were the additional variables that differed significantly among major teaching fields.

L scores of 0, 1, and 2 correspond approximately to the Rorschach W (whole), D (large detail), Dd (small detail) scores. In Rorschach interpretation, emphasis is placed on the percentage relationship of W, D and Dd. As such, the Holtzman pooled L score is not directly comparable to the W, D, and Dd but does suggest perceptual tendencies. In the Rorschach, a high W percentage score usually is a sign of organizational interest and ability. The subject possessing such ability is interested in and capable of viewing the relatively separate facets of his experience as an interrelated whole. A high D percentage score suggests interest and ability to differentiate

perceptually, with relatively little interest in integration and organization. This could be interpreted as a practical, everyday, common-sense application of intelligence, an interest in the presented, obvious facts without much drive to seek relationships between presented facts of experience. A high Dd percentage score in the Rorschach is obtained by selecting for the response a small blot area and may represent not only a differentiated interest in factual things, but often is associated with an emphasis on accuracy, correctness and exactness or sometimes simply reflects a need for certainty or minutiae of life. Thus, it would appear that the higher the Holtzman L score, the more the subject perceives the minutiae of life and, conversely, the lower the L score, the more apparent is the tendency to view life as interrelated.

There is nothing in the traditional Rorschach scoring system which approximates the Holtzman Br variable. At best, the interpretation of a high Br score would suggest an interest in being protected from danger, harm, or the unknown, suggesting an underlying fear of some person, place, act or animal.

Generally, the rationale of the use of shading is based on the general hypothesis that the way in which a person handles the shading aspects of a blot is related to the way the subject handles his primary security need

and derived needs for affection and belongingness. Since some frustration of affectional need is practically universal and all persons have more or less anxiety from this source, the mere presence of shading in a response does not suggest anything out of the ordinary. However, the higher the Sh variable score, the more probable is the hypothesis that there is a significant amount of anxiety and frustration over affectional needs present.

Hypotheses III,A and III,B

As previously discussed, the investigation of hypothesis II identified several variable scores which significantly differed between the sexes and it was therefore logical to re-examine hypothesis III by grouping the sample according to sex prior to investigating for significant differences among student teachers grouped according to major teaching fields.

Hypothesis III,A stated that there would be no significant differences among each of the mean Holtzman variable scores of MALE student teachers grouped according to major teaching field; hypothesis III,B was similar but for FEMALE student teachers.

The test of hypothesis III,A shows that of the seventeen variables only one, At (anatomy), reached statistical significance at the five per cent level. The At variable is a content score and as such presents

interpretive difficulties discussed previously on page 124. The same comments would apply here and would suggest the same speculative interpretation with no conclusive interpretive remarks.

The investigation of hypothesis III,B, however, revealed that ten variable mean scores (L, location; R, rejection; Sh, shading; M. movement; I, integration; A, animal; Sx, sex; Hs, hostility; Br, barrier; and P, popular) are significantly different among female student teachers grouped according to major teaching field. Six of these variable scores (R, rejection; M, movement; I, integration; A, animal; Sx, sex; and P, popular) were also significantly different between the sexes (Table 10) and discussed on pages 118 through 125 and the remaining variables of L (location), Sh (shading), and Br (barrier) were significantly different among major teaching fields and discussed on pages 126-128.

The remaining significantly different variable, Hs (hostility) of females grouped according to major teaching field, was not significantly different in any other hypothesis.

The Hs or Hostility variable concept score represents symbolic, implicit or explicit signs of hostility. Although the mean Hs score does not differentiate between the sexes of the student teachers, females, when considered as a group score significantly different

among each other when classified according to major teaching field. Why this variable is significantly different, or what interpretations can be given to this finding is unknown. And other than reporting the fact, little else can be said with any degree of certitude.

Perhaps it is appropriate, in summary, to again call the reader's attention to the purpose of the study. Prior to the undertaking of this investigation, no relationships of Holtzman variables to success in student teaching or major teaching fields were known to exist. At best, the study could have identified significant differences in such a way as to suggest group characteristics which could be studied in some later, more individual, or clinical inquiries. Failing to find clear significant differences in the variable scores when the sample was grouped according to various classifications did provide answers to the hypotheses posed and added more information which might be useful to other investigators and further studies in this area.

CHAPTER IV

SUMMARY, CONCLUSIONS AND RECOMMENDATIONS

Summary

The present research studied the relationship between computer scored group Holtzman Inkblot variable mean scores and student teaching success (Table 9). The study also investigated the relationship between variable scores of student teachers grouped according to sex (Table 10) and grade/subject teaching fields (Table 14). The study also generated norms (Table 8) for student teachers on the group Holtzman Inkblot Test and determined the feasibility of administering and computer scoring a projective technique with a large sample.

The group version of the Holtzman Inkblot Test was administered to a total of 802 Michigan State University student teachers during the Winter Term, 1968. The test was administered in each of the geographical centers where Michigan State places student teachers. Forty-five inkblots were exposed to each group for one minute each and the student teachers responded with a short written description of what each inkblot "looked like" on the specially designed group answer sheet (Appendix D). The answer sheets were alpha-numeric coded to identify the

sex and teaching field/grade level of the student teachers. After collecting all responses from the student teachers, the data and responses were keypunched according to the format specified by Gorham[1] (Appendix E), loaded onto magnetic tape and, through special arrangements with Gorham, the tape was submitted for computer scoring to the Holtzman Inkblot Scoring Service in Perry Point, Maryland. Output from the scoring service included a print-out and a score summary card (Appendix E) for each student teacher on seventeen Holtzman variables: Location, Rejection, Form Definiteness, Color, Shading, Movement, Integration, Human, Animal, Anatomy, Sex, Abstract, Anxiety, Hostility, Barrier, Penetration and Popular.

The study was limited to testing the following hypotheses: (1) Hypotheses I, I,A and I,B stated that there would be no significant difference on the Holtzman variable mean scores among (I--all; I,A--male; I,B--female) student teachers classified according to rated success by the university coordinators. (2) Hypothesis II stated that there would be no significant differences on the Holtzman variable mean scores when the student teachers were classified according to sex, and

---

[1]Donald R. Gorham, The Development of a Computer Scoring System for Inkblot Responses, Report #1, Psychological Research Reports, Research Laboratory, Psychology Service, Perry Point, Maryland, September, 1965.

(3) Hypothesis III stated that there would be no significant differences on the variable scores when (III--all; III,A--male; III,B--female) student teachers were grouped according to teaching field or grade level.

Conclusions

Tables 9, 10, 11, 12, 13, 14 and 15 show the computed means, standard deviations, F statistic and its approximate significance for testing each null hypothesis. The analyzed data produced the following findings:

Hypothesis I--Conclusions

The statistical tests performed on hypothesis I show that, with the exception of one variable (R, rejection), there are no significant differences among the mean Holtzman variable scores of student teachers grouped according to rated success. Thus, the conclusion is that the Holtzman scores are of little or no value in discriminating among the success ratings of the student teachers.

Hypothesis I,A and I,B--Conclusions

The statistical tests performed on hypotheses I,A and I,B show that, with the exception of one variable (R, rejection) there are no significant differences among the mean Holtzman variable scores of male (I,A) or female (I,B) student teachers grouped according to rated success.

Thus, the conclusion is that the Holtzman scores are of little value in discriminating among the success ratings of either male or female student teachers.

Hypothesis II--Conclusions

The statistics for testing hypothesis II show that eleven of the seventeen computer scored Holtzman variable mean scores significantly differentiate between student teachers grouped according to sex. Thus, the conclusion is that the Holtzman scores are of some but restricted value in discriminating between student teachers grouped according to sex.

Hypothesis III--Conclusions

When the sample of student teachers (excluding those at the junior high school teaching level) were grouped into thirteen major teaching areas, the test of hypothesis III showed that nine of the seventeen Holtzman variable mean scores differentiate among the teaching areas. Thus, the conclusion is that the Holtzman scores are of some but restricted value in discriminating among student teachers grouped according to major teaching field.

Hypothesis III,A and III,B--Conclusions

The statistical tests performed on hypotheses III,A and III,B show that when the student teachers are first grouped according to sex (male, III,A; female, III,B) and

then teaching field, there are no significant differences among the mean variable scores, with the exception of one variable, At (anatomy), for the male student teachers. Thus, the conclusion is that the Holtzman variable scores are of little or no value in discriminating among male student teachers grouped according to major teaching field.

For females, however, grouped according to major teaching field (hypothesis III,B), eleven of the seventeen mean variable scores are significantly different. Thus, the conclusion is that the Holtzman variable scores are of some but restricted value in discriminating among female student teachers grouped according to major teaching field.

Discussion

The conclusions resulting from the statistical tests of the hypotheses are not claimed for any group other than the student teachers represented in this study. The data for the study were obtained entirely at Michigan State University during the Winter Term of 1968 and similar conclusions for other groups at other times cannot be inferred from this particular study.

Perhaps as important, though, as statements regarding the generality of the conclusions, are comments on the assumptions and framework within which the study was conceived.

If, as Combs[2] suggests, "whether an individual will be an effective teacher depends upon the nature of his private world of perceptions," then either this statement is an oversimplification of a highly complex reality or, insofar as the Holtzman Test measures perception, it is without supportive evidence at this point. However, even this may not be quite so.

It was assumed in this study that personality is an important variable in successful teaching. This is well documented in the general literature and common sense suggests this is the case, but objective, empirical evidence is lacking. The problem stems, in part, from difficulties in personality definition, criteria and measurement. And the problem is exceedingly compounded by the relationship between reported perception and personality, not to mention difficulties in criteria and measurement of successful teaching. If projective techniques measure perception, as most authorities would agree, their meaning is highly clinical or individual in nature and only the most general statements can be made on the basis of a group projective technique. A concept of a "group personality" is a contradiction in terms since personality as such has its meaning only in an individual sense.

---

[2]Combs, The Professional Education of Teachers, p. 19.

And then, too, it was assumed that the university coordinator could judge success in student teaching. It would be unwarranted to believe that subjective judgments of success, even with specific criteria, are in any way objective. Yet, the present state of teaching success research suggests that supervisory ratings are the most appropriate measure of success. Concepts of teaching success vary from one judge to another, and in this study the variance of success rating among coordinators is unspecified. What is success in student teaching as judged by one coordinator may not agree with another's judgment. Perhaps this is why no significant differences in mean Holtzman scores were found when student teachers were grouped according to rated success. In other words, perhaps the assumption that the coordinators' rating of success on the basis of performance discriminates among student teachers who were more successful than others in their experience is a questionable assumption and clouds over the discovery of significant mean score differences.

The problem of rating success in not an easy one, but it cannot be simply ignored. It is caused by the absence of any generally agreed upon criteria of teaching success and objective methods of measurement. This study, or any similar study, would be dramatically improved by specification of teaching success criteria and objective

measurement, thus eliminating the subjective element in success judgments.

In sum, it may be that the primary obstacle in discovering distinctive and specific features of the successful student teacher personality is due generally to the lack of an adequate theory of teaching and its resultant lack of success criteria, subjectivity in judgment even when the criteria are specified, problems in relating perceptual modes to individual personality from a group test, and the effect of subjectivity in all of these cases.

In any event, the feasibility of such an undertaking has been evaluated. That is, it is technically and feasibly possible to administer and score a group projective technique to a large sample of subjects. The technique of computer scoring permits an endless variety of individual responses to be scored objectively (that is, the only subjective element is that which is built into the computer scoring program and which is a constant for the scoring of all protocols) with a minimum amount of time and effort. However, the interpretation of results should only be on an individual basis since a "group" interpretation is meaningless and only serves to describe the group in a most general sense. From the data collected, it would appear that the classification of all responses of all the subjects into seventeen

categories does not produce common perceptual modes. That is, the private world of perceptions, as measured by the Holtzman Inkblot variables, of successful student teachers does not differ significantly from the private world of perceptions of less successful student teachers.

Recommendations

From the presentation, analysis, and discussion of the data, the following recommendations are made:

1. A longitudinal study of the student teachers should be undertaken to follow-up their teaching success as a first year teacher. This would permit a determination of the relationship between rated student teaching success and success on the job.

2. A study should be made to determine the variance in success judgments among university coordinators. An ideal study in this area would have all coordinators judge the success of a sample of student teachers and in this fashion determine the amount of variance, if any, that exists in coordinators' success judgments. If a significant amount of variance does exist, there would be cause to re-evaluate the data in this, and other, similar studies.

3. It is also recommended that an in-depth study be made of the university coordinator to determine if personality is a significant factor in the university coordinator's evaluation of the student teacher.

4. It is recommended that the significance of the R (rejection) score when related to rated success be pursued with specific measures of "openness" or "psychological defensiveness" and its relationship to success in student teaching.

5. There should be a continued search for those personality characteristics related to successful student teaching. This should be done, however, by clinical or individual methods involving those student teachers at the extreme ends of the rating scale and comparing the individual protocols for systematic differences or similarities.

6. It is recommended that studies in the area of personality and student teaching success be less global in nature and utilize projective techniques that introduce an interaction or sociological dimension such as might be found in a teacher-type Thematic Apperception Test.

7. It is suggested that since the sexes scored significantly different on nine variables, future studies should continue to analyze differences in light of these findings and relate the data to cultural stereotypes of male and female temperaments.

8. It is also recommended that since females, when grouped according to major teaching area, scored significantly different on eleven variables, future studies

should analyze differences in the light of major teaching field and relate the data to teacher stereotypes of subject matter areas.

9. Finally, it is recommended that further work investigating personality and teaching success be undertaken when there is general agreement, either institutional or profession-wide, on what is successful teaching, derived from an overall theory of teaching.

BIBLIOGRAPHY

BIBLIOGRAPHY

Abt, Lawrence E., and Bellack, Leopold. Projective Psychology. New York: Alfred A. Knopf, 1950.

Alexander, T. "The Adult-Child Interaction Test: A Projective Test for Use in Research." Monograph Social Research in Child Development, Vol. XVII (No. 2, 1952).

Allen, Thomas K. "Personality Patterns of Student Teachers Planning to Teach Different Subjects." Journal of the Student Personnel Association for Teacher Education, Vol. IV (December, 1965), 89-93.

_____. "Personality as a Predictor of Student Teaching Success." Journal of the Student Personnel Association for Teacher Education, Vol. VI (Fall, 1966), 12-16.

Allport, Gordon W. Pattern and Growth in Personality. New York: Holt, Rinehart and Winston, 1961.

Barr, Arvil S. "The Measurement and Prediction of Teacher Efficiency, A Summary of Investigations." Journal of Experimental Education, Vol. XVI (June, 1948), 203-283.

_____. "Pupil Changes and Teacher Ability," Journal of Experimental Education, Vol. XIV (September, 1955), 52-74.

_____. "Wisconsin Studies of the Measurement and Prediction of Teacher Effectiveness: A Summary of Investigations." Journal of Experimental Education, Vol. XXX (September, 1961), 5-156.

Biddle, Bruce J., and Ellena, William J. Contemporary Research on Teacher Effectiveness. New York: Holt, Rinehart and Winston, 1964.

Charles, Harvey. "The Use of a Projective Technique in Teacher Selection." National Catholic Educational Association Bulletin, Vol. LVII (August, 1961), 172-173.

Clark, E. J. "The Mental Health of Elementary School Teachers as Measured by the Guilford-Martin Personality Battery." Paper read at the National Council on Measurements Used in Education, Atlantic City, March, 1950.

Cohan, Richard W. "The Holtzman Inkblot Test." The Sixth Mental Measurements Yearbook. Edited by O. K. Buros. 3rd ed. Highland Park, New Jersey: The Gryphon Press, 1965.

Combs, Arthur W. "The Personal Approach to Good Teaching." Educational Leadership, Vol. XXI (March, 1964), 269-377.

_____. The Professional Education of Teachers. Boston: Allyn and Bacon, Inc., 1965.

Cook, Desmond, et al. "A Factor Analysis of Teacher Trainee Responses to Selected Personality Inventories." Educational and Psychological Measurement, Vol. XXI (Winter, 1961), 865-872.

Cronbach, L. J. "Statistical Methods Applied to Rorschach Scores: A Review." Psychological Bulletin, Vol. XXXXVI (September, 1949), 393-429.

Dixon, W. R., and Morse, W. C. "The Prediction of Teaching Performance: Empathic Potential." Journal of Teacher Education, Vol. XII (September, 1961), 322-329.

Domas, Simeon J., and Tideman, David V. "Teacher Competence: An Annotated Bibliography." Journal of Experimental Education, Vol. XIX (December, 1950), 101-218.

Erickson, Harley E. "A Factorial Study of Teaching Ability." Journal of Experimental Education, Vol. XXIII (September, 1954), 1-39.

Evans, Kathleen M. "A Critical Survey of Methods of Assessing Teacher Ability." British Journal of Educational Psychology, Vol. XXI (June, 1951), 89-95.

Fattu, Nicholas A. "Effectiveness--An Elusive Quality." The Education Digest, Vol. XXVII (January, 1962), 24-26.

Flanagan, Carroll E. "A Study of the Relationship of Scores of the MMPI to Success in Teaching as Indicated by Supervisory Ratings." Journal of Experimental Education, Vol. XXXIX (June, 1961), 329-354.

Frank, L. K. "Projective Methods for the Study of Personality." Journal of Psychology, Vol. VII (1939), 389-413. Quoted in Klopfer, Bruno, and Davidson, Helen H. The Rorschach Technique: An Introductory Manual. New York: Harcourt, Brace and World, Inc., 1962.

Gage, N. L. Address appearing in "Proceedings." Research Resume. Burlingame, California: California Teachers Association, 1960.

Getzels, J. W., and Jackson, P. W. "The Teacher's Personality and Characteristics." Handbook of Research on Teaching. Edited by N. L. Gage. Chicago: Rand McNally and Company, 1963.

Goodenough, Florence L.; Fuller, Elizabeth M.; and Olson, Edna. "The Use of the Goodenough Speed of Association Test in the Pre-Service Selection of Nursery School-Kindergarten-Primary Teachers." Journal of Educational Psychology, Vol. XXXVII (September, 1946), 335-346.

Gorham, Donald R. The Development of a Computer Scoring System for Inkblot Responses. Report #1, Psychological Research Reports, Research Laboratory, Psychology Service, Perry Point, Maryland, September, 1965.

_____. "Validity and Reliability of a Computer Based Scoring System for Inkblot Responses." Journal of Consulting Psychology, Vol. XXXI (February, 1967), 65-70.

Gorham, Donald R.; Moseley, Edward C.; and Holtzman, Wayne H. "Norms for Computer-Scored Holtzman Inkblot Technique." Perceptual and Motor Skills, Vol. XXVI (June, 1968), 1279-1305.

Gowan, J. C., and Gowan, May S. "The Guilford-Zimmerman and the California Psychological Inventory in the Measurement of Teacher Candidates." California Journal of Educational Research, Vol. VI (January, 1955), 35-37.

Gustafson, Kent L. "Simulation of Anxiety Situations and Its Resultant Effect on Anxiety and Classroom Interaction of Student Teachers." Unpublished Ph.D. dissertation, Michigan State University, 1969.

Hadley, S. T. "A Study of the Predictive Value of Several Variables of Student Teaching Success as Measured by Student Teaching Marks." Teacher's College Bulletin, Vol. LX (No. 3, 1954), 1-10.

Hamacheck, Donald. "Characteristics of Good Teachers and Implications for Teacher Education." Phi Delta Kappan, Vol. L, No. 6 (February, 1969), 341-344.

Hampton, N. D. "An Analysis of Supervisory Ratings of Elementary Teachers Graduated from Iowa State Teachers College." Journal of Experimental Education, Vol. XX (December, 1951), 180-215.

Hathaway, S. R., and McKinley, J. C. Minnesota Multiphasic Personality Inventory Manual. New York: The Psychological Corporation, 1951.

Hertz, Marguerite R. "The Use and Misuse of the Rorschach Method." Journal of Projective Techniques, Vol. XXIII (March, 1959), 33-48.

Hicks, William Vernon, and Blackington, Frank H. Introduction to Education. Columbus, Ohio: Charles E. Merrill Books, Inc., 1965.

Hilton, Thomas A. "Ego Involvement in Teaching: Its Theory and Measurement by a Word Completion Technique." Unpublished Ph.D. dissertation, Harvard University, 1955.

Holtzman, Wayne H., et al. Inkblot Perception and Personality. Austin, Texas: University of Texas Press, 1961.

_____. "Comparison of the Group Method and the Standard Individual Version of the Holtzman Inkblot Technique." Journal of Clinical Psychology, Vol. XIX (October, 1963), 441-449.

Holtzman, Wayne H. "The Holtzman Inkblot Technique." Projective Techniques in Personality Assessment. Edited by Albert Rabin. New York: Springer Publishing Company, 1968.

Inlow, Gail M. "Comparative Study of Student Teaching Practices in 38 Midwestern Institutions." Journal of Experimental Education, Vol. XXVIII (June, 1960), 337-349.

Jones, Margaret L. "Analysis of Certain Aspects of Teaching Ability." Journal of Experimental Education, Vol. XXV (December, 1956), 152-180.

Kemp, L. D. "The Prediction of Teaching Success." The Forum of Education, Vol. VI (August, 1947), 14.

Klopfer, Bruno, and Davidson, Helen H. The Rorschach Technique: An Introductory Manual. New York: Harcourt, Brace and World, Inc., 1962.

Lamke, Thomas A. "Personality and Teaching Success." Journal of Experimental Education, Vol. XX (December, 1951), 217-259.

Leeds, C. H. "Teacher Attitude and Temperament as a Measure of Teacher-Pupil Rapport." Journal of Applied Psychology, Vol. XXXX (October, 1956), 333-337.

LePere, Jean M., and Cox, Richard C. Training Elementary Teachers: Comparisons of Separate and Block Methods Courses. East Lansing, Michigan: Bureau of Educational Research Services, Michigan State University, 1964.

Lewis, James N. "The Relationship of Attributes Measured by the Structured-Objective Rorschach Test and Success in Student Teaching." Unpublished Ed.D. dissertation, North Texas State University, 1966.

Lindsey, Margaret. New Horizons for the Teaching Profession. Washington, D.C.: National Commission on Teacher Education and Professional Standards, National Education Association, 1961.

Lough, Orpha M. "Women Students in Liberal Arts, Nursing, and Teacher Training Curricula and the MMPI." Journal of Applied Psychology, Vol. XXXI (August, 1947), 437-445.

Lynch, W. W. "Person Perception: Its Role in Teaching." Indiana University School of Education Bulletin, Vol. XXXVII (1962), 1-37.

Magee, Robert M. "Admission-Retention in Teacher Education." Journal of Teacher Education, Vol. XII (March, 1961), 81-85.

McClendon, P. E. "Teacher Perception and Working Climate." Educational Leadership, Vol. XX (December, 1962), 104-108.

Mitzel, Harold E. "Criteria of Teacher Effectiveness." Encyclopedia of Educational Research. Edited by C. W. Harris. New York: MacMillian and Company, 1960.

Montross, Harold W. "Temperament and Teaching Success." Journal of Experimental Education, Vol. XXIII (September, 1954), 34-97.

Moore, Clark H., and Cole, David. "The Relation of MMPI Scores to Practice Teaching Ratings." Journal of Educational Research, Vol. L (May, 1957), 711-716.

Morsh, Joseph H., and Wilder, Eleanor W. Identifying the Effective Instructor: A Review of the Quantitative Studies, 1900-1952 (Research Bulletin, AFPTRC - TR 54:44). San Antonio: United States Air Force Personnel Training Research Center, 1954.

National Commission on Teacher Education and Professional Standards. The Education of Teachers: Considerations in Planning Institutional Programs. Washington, D.C.: National Education Association, 1960.

Ohlsen, Merle M., and Schultz, Raymond E. "Projective Test Response Patterns for Best and Poorest Student Teachers." Education and Psychological Measurement, Vol. XV (Spring, 1955), 18-27.

Osmon, Robert V. The Improvement of Secondary Teaching. St. Louis: Education Publishers, 1962.

Pettinger, B. F. "Problems of Teacher Measurement." Journal of Educational Psychology, Vol. VIII (January, 1917), 103-110.

Phillips, L; Kaden, S.; and Waldman, M. "Rorschach Indices of Developmental Level." Journal of Genetic Psychology, Vol. LXXXXIV (1959), 267-285.

Rabin, Albert I. Varieties of Rorschach Interpretation. Paper delivered in symposium on "The Uses and Misuses of the Rorschach," at the Michigan Psychological Association meetings in Detroit, May, 1958.

Ramsey, C. P. "Leadership Perception." Educational Leadership, Vol. XX (December, 1962), 151-154.

Richter, H., and Winter, W. D. "Holtzman Inkblot Correlates of Creative Potential." Journal of Projective Techniques and Personality Assessment, Vol. XXX (February, 1966), 62-67.

Rolfe, J. F. "The Measurement of Teaching Ability." Journal of Experimental Education, Vol. XIV (September, 1945), 52-74.

Ryans, David G. "The Criteria of Teaching Effectiveness." Journal of Educational Research, Vol. XLII (May, 1949), 690-699.

_____. "Prediction of Teacher Effectiveness." Encyclopedia of Educational Research. Edited by C. W. Harris. 3rd ed. New York: MacMillian Company, 1960.

_____. Characteristics of Teachers. Washington, D.C.: American Council on Education, 1960.

_____. "Assessment of Teacher Behavior and Instruction." Review of Educational Research, Vol. XXXIII (October, 1963), 415-1441.

Steeves, Frank L. "Crucial Issues in Student Teaching." Journal of Teacher Education, Vol. XVI (September, 1965), 307-310.

Stevens, S. S. "Psychology and the Science of Science." Psychological Bulletin, Vol. XXXVI (April, 1939), 221-263.

Stout, Ruth A. "Practices for Selection in Teacher Education." Teacher Education: The Decade Ahead. Washington, D.C.: National Commission on Teacher Education and Professional Standards, National Education Association, 1955.

_____. "Selective Admissions and Retention Practices in Teacher Education." Journal of Teacher Education, Vol. VIII (December, 1957), 422-432.

Swartz, Jon D., and Holtzman, Wayne H. "Group Method of Administration for the Holtzman Inkblot Technique." Journal of Clinical Psychology, Vol. XIX (October, 1963), 433-441.

Sultan, E. E. "A Quantitative Investigation of the Rorschach Inkblot Test as Applied to Student Teachers." British Journal of Social and Clinical Psychology, Vol. IV (September, 1965), 197-206.

Symonds, Percival M. "Characteristics of the Effective Teacher Based on Pupil Evaluations." Journal of Experimental Education, Vol. XXIII (June, 1955), 289-310.

Symonds, Percival M., and Dudek, Stephanie. "Use of the Rorschach in the Diagnosis of Teacher Effectiveness." Journal of Projective Techniques, Vol. XX (June, 1956), 227-234.

Tarpey, Sister M. S. "Personality Factors in Teacher Trainee Selection." British Journal of Educational Psychology, Vol. XXXV (June, 1965), 140-149.

Thetford, William H. "The Holtzman Inkblot Technique." The Sixth Mental Measurements Yearbook. Edited by O. K. Buros. 3rd ed. Highland Park, New Jersey: The Gryphon Press, 1965.

Underhill, Robert G. "The Relation of Elementary Student Teacher Empathy (Affective Sensitivity) Change to Supervising Teacher Empathy and Student Teaching Success." Unpublished Ed.D. dissertation, Michigan State University, 1968.

Willcox, Isabel, and Beigel, Hugo G. "Motivations in the Choice of Teaching." Journal of Teacher Education, Vol. IV (June, 1953), 106-109.

Witty, Paul. "Some Characteristics of Effective Teachers." Educational Administration and Supervision, Vol. XXXVI (April, 1950), 193-208.

Zubin, J. "Failure of the Rorschach Technique." Journal of Projective Techniques, Vol. XVIII (September, 1954), 303-315.

APPENDICES

APPENDIX A

DESCRIPTION OF THE 17 COMPUTER-SCORED HIT VARIABLES

DESCRIPTION OF THE 17 COMPUTER-SCORED HIT VARIABLES

Variable	Brief Description	Score Range
L	A 3-point scale* used to measure the tendency to fragment the inkblot into smaller areas. The greater the area used, the lower the score.	0-90
R	Failure of *S* to give a scorable response.	0-45
FD	A 5-point scale measuring the definiteness of the form of the concept reported.	0-180
C	A 4-point scale measuring the primacy of both chromatic and achromatic color in determining a response.	0-135
Sh	A 3-point scale used to measure the primacy of shading as a response determinant.	0-90
M	A 5-point scale indicating the movement energy level *S* voluntarily ascribes to his percept regardless of its content.	0-180
I	A 2-point scale indicating the presence or absence of the organization of adquately perceived inkblot elements into a unified response.	0-45
H	A 3-point scale used to measure the amount of human content seen. Score 1 is recorded for human details, and 2 for a whole human.	0-90
A	A 3-point scale used to measure the amount of animal content seen. Score 1 is recorded for animal details and 2 for a whole animal.	0-90
At	A 3-point scale used to measure the amount of anatomy content seen. Score 1 is recorded for bone structures, and 2 for visceral and crude anatomy responses.	0-90
Sx	A 3-point scale used to measure the amount of sexual content seen. Score 1 is recorded for socially acceptable sexual content and 2 for more blatant sexual responses.	0-90
Ab	A 3-point scale characterized by relative or total independence of the response from the formal blot qualities. For instance; "happiness," "hope."	0-90
Ax	A 3-point scale used to measure the amount of anxiety reflected by the response.	0-90
Hs	A 4-point scale used to measure the amount of hostility elicited by the blot stimulus.	0-135
Br	A 2-point scale used to measure the presence or absence of body image boundary in a response.	0-45
Pn	A 2-point scale used to measure the presence or absence of body image boundary diffusion or disruption in a response.	0-45
P	A 2-point scale used to measure the presence or absence of a popular content. A response is judged popular which occurs at least once in seven protocols for normal *Ss*.	0-45

*In all other scales except *R*, the zero point represents an absence of the characteristic being measured.

APPENDIX B

SUCCESS RATING CHART

SUCCESS RATING CHART

CRITERIA OF SUCCESS

A. Working with People

B. Establishing Classroom Climate

C. Planning Instruction

D. Managing Instruction

E. Command of Subject and Teaching Materials

F. Personal Qualities

G. Professional Qualities

H. General Effectiveness as a Teacher

SUCCESS RATING NUMBER	INTERPRETATION	PERCENTAGE OF STUDENTS LIKELY TO RECEIVE THIS NUMBER
1	**ONE OF THE VERY BEST STUDENT TEACHERS I HAVE EVER SEEN.** Assign rating 1 to the MOST EXCEPTIONAL and OUTSTANDING student teacher of all. If you judge that a student is of ABSOLUTELY OUTSTANDING ACCOMPLISHMENT and will make a potentially GREAT and PROFOUND effect on students, assign rating 1.	2% or about 15 out of 850
2	**HIGHLY SUCCESSFUL.** Assign rating 2 to those of somewhat lesser overall ability than rating 1 but nevertheless represent ACCOMPLISHED and OUTSTANDING student teaching performance. The highly successful student would rate close to rating 1 but is not one of the best student teachers you have ever seen.	13% or about 110 out of 850
3	**SUCCESSFUL.** Assign rating 3 to COMPETENT student teachers. Most should receive this rating and do not show the outstanding qualities of ratings 1 or 2.	About 70% or most or about 595 out of 850
4	**LESS SUCCESSFUL** Assign rating 4 to those student teachers who have some problems and rate below the middle, that is, competent and successful, group in your center.	13% or about 110 out of 850
5	**PASSED BUT SHOULD NOT BE IN TEACHING.** Assign rating 5 to those students who you feel OUGHT to fail - they really are not inclined to teaching - but because of various reasons should not receive a failing grade.	2% or about 15 out of 850
6	**FAILED.** Assign rating 6 to all student teachers who fail student teaching in your center.	
7	**DROPPED/GRADE POSTPONED.** Assign rating 7 to all student teachers who DROPPED student teaching or had their GRADE POSTPONED for any reason at all.	

APPENDIX C

COVER LETTER TO STUDENT TEACHERS

MICHIGAN STATE UNIVERSITY EAST LANSING · MICHIGAN 48823

COLLEGE OF EDUCATION · STUDENT TEACHING OFFICE · 253 ERICKSON HALL

January, 1968

Dear M.S.U. Student Teacher,

I am a graduate student doing research on the perceptions of student teachers in connection with the Student Teaching Office at Michigan State University.

This study is concerned with an investigation of the relationships between perceptual modes, chosen teaching field and grade level. I will be administering a Rorschach-like inkblot method to all Winter Term student teachers.

You are one of the student teachers who can help establish perceptual norms for your teaching field and grade level. I have contacted your student teaching coordinator to arrange details of when and where the inkblots will be shown in your center. And your coordinator will announce this meeting time for your center group.

Let me assure you that all of the information will be confidential. The results will be summarized only in tables, percentages, and so on.

A brief report of the findings will be available in the Student Teaching Office on campus after April, 1968.

Your cooperation is sincerely appreciated.

Very truly yours,

B. Bradley West
Administrative Assistant
Student Teaching Office

APPENDIX D

HIT RESPONSE SHEET

HIT RESPONSE SHEET

2/3

Teaching #

NAME Last First Student number

4/5/6

Subject #

7

A

Form

HOME ADDRESS Street City State

8/9

Age (Yrs.)

DATE Month / Day / Year Marital status

JR SR GRAD SPEC

CLASS (Circle) Approx. G.P.A.

10

M F

Sex (circle)

11/12 13

	0	1	2	
01	1	½	¼	
02	1	½	¼	
03	1	½	¼	
04	1	½	¼	
05	1	½	¼	
06	1	½	¼	
07	1	½	¼	
08	1	½	¼	
09	1	½	¼	
10	1	½	¼	
11	1	½	¼	
12	1	½	¼	
13	1	½	¼	
14	1	½	¼	
15	1	½	¼	
16	1	½	¼	
17	1	½	¼	
18	1	½	¼	
19	1	½	¼	
20	1	½	¼	

21	1 ½ ¼	
22	1 ½ ¼	
23	1 ½ ¼	
24	1 ½ ¼	
25	1 ½ ¼	
26	1 ½ ¼	
27	1 ½ ¼	
28	1 ½ ¼	
29	1 ½ ¼	
30	1 ½ ¼	
31	1 ½ ¼	
32	1 ½ ¼	
33	1 ½ ¼	
34	1 ½ ¼	
35	1 ½ ¼	
36	1 ½ ¼	
37	1 ½ ¼	
38	1 ½ ¼	
39	1 ½ ¼	
40	1 ½ ¼	
41	1 ½ ¼	
42	1 ½ ¼	
43	1 ½ ¼	
44	1 ½ ¼	
45	1 ½ ¼	

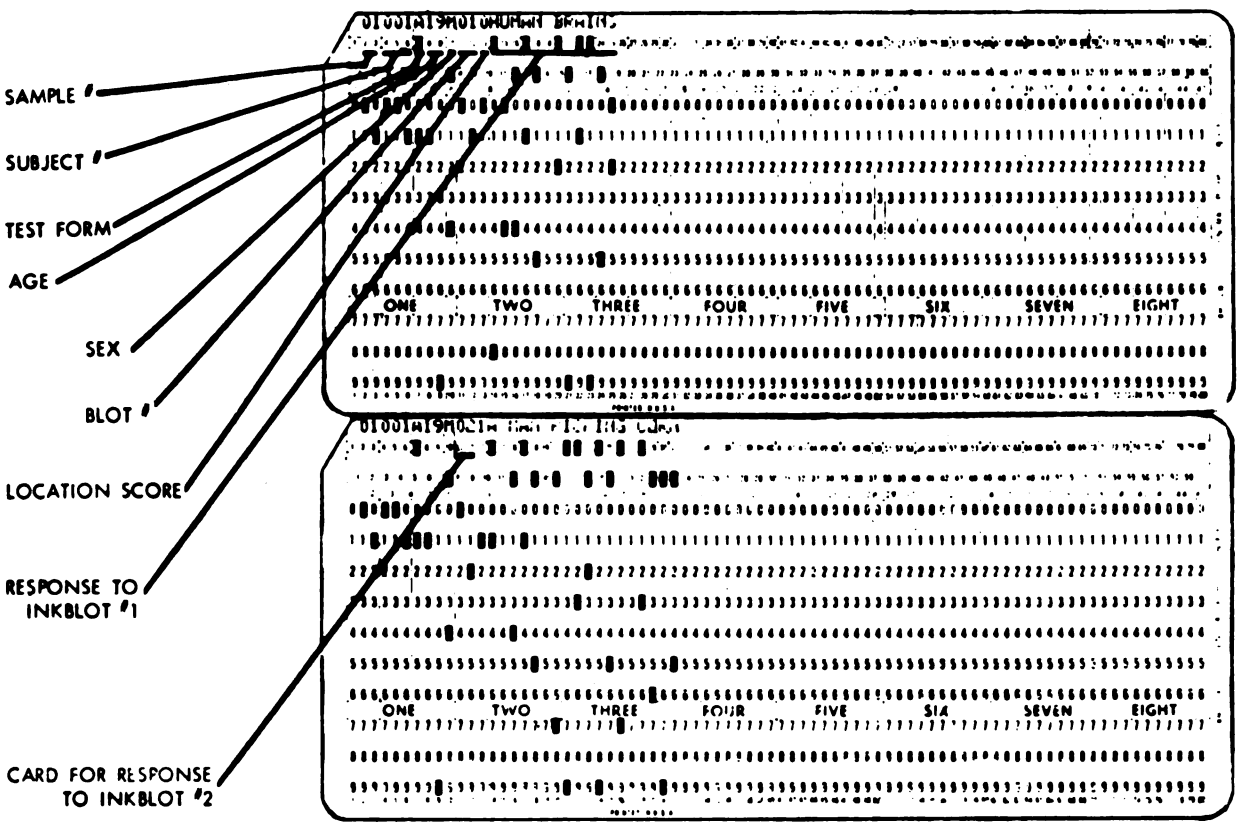
21	1 $\frac{1}{2}$ $\frac{1}{4}$	
22	1 $\frac{1}{2}$ $\frac{1}{4}$	
23	1 $\frac{1}{2}$ $\frac{1}{4}$	
24	1 $\frac{1}{2}$ $\frac{1}{4}$	
25	1 $\frac{1}{2}$ $\frac{1}{4}$	
26	1 $\frac{1}{2}$ $\frac{1}{4}$	
27	1 $\frac{1}{2}$ $\frac{1}{4}$	
28	1 $\frac{1}{2}$ $\frac{1}{4}$	
29	1 $\frac{1}{2}$ $\frac{1}{4}$	
30	1 $\frac{1}{2}$ $\frac{1}{4}$	
31	1 $\frac{1}{2}$ $\frac{1}{4}$	
32	1 $\frac{1}{2}$ $\frac{1}{4}$	
33	1 $\frac{1}{2}$ $\frac{1}{4}$	
34	1 $\frac{1}{2}$ $\frac{1}{4}$	
35	1 $\frac{1}{2}$ $\frac{1}{4}$	
36	1 $\frac{1}{2}$ $\frac{1}{4}$	
37	1 $\frac{1}{2}$ $\frac{1}{4}$	
38	1 $\frac{1}{2}$ $\frac{1}{4}$	
39	1 $\frac{1}{2}$ $\frac{1}{4}$	
40	1 $\frac{1}{2}$ $\frac{1}{4}$	
41	1 $\frac{1}{2}$ $\frac{1}{4}$	
42	1 $\frac{1}{2}$ $\frac{1}{4}$	
43	1 $\frac{1}{2}$ $\frac{1}{4}$	
44	1 $\frac{1}{2}$ $\frac{1}{4}$	
45	1 $\frac{1}{2}$ $\frac{1}{4}$	

APPENDIX E

SAMPLE OF INPUT

SAMPLE OF OUTPUT

SAMPLE OF INPUT



SAMPLE OF OUTPUT

TEST SCORING RUN OF 12 CASES FROM SAMPLE 01

SAMPLE 01 SUBJECT 001 TEST FORM A SEX=MALE AGE=19

SCORES	NH	L	R	FC	C	SH	M	I	H	A	AT	SX	AB	AX	MS	BR	PN	P
	45	30	2	67	14	7	13	2	15	23	8	0	0	1	2	10	1	10

SAMPLE 01 SUBJECT 002 TEST FORM A SEX=FEMALE AGE=17

SCORES	NH	L	R	FC	C	SH	M	I	H	A	AT	SX	AB	AX	MS	BR	PN	P
	43	50	0	91	3	1	32	3	26	34	2	0	1	15	11	8	4	11

SAMPLE 01 SUBJECT 003 TEST FORM A SEX=MALE AGE=21

SCORES	NH	L	R	FC	C	SH	M	I	H	A	AT	SX	AB	AX	MS	BR	PN	P
	43	42	1	86	5	1	21	4	23	32	5	0	1	9	14	14	7	10

SAMPLE 01 SUBJECT 004 TEST FORM A SEX=MALE AGE=20

SCORES	NH	L	R	FC	C	SH	M	I	H	A	AT	SX	AB	AX	MS	BR	PN	P
	43	24	5	60	11	11	31	4	17	17	2	2	2	5	6	6	4	11

SAMPLE 01 SUBJECT 005 TEST FORM A SEX=MALE AGE=20

SCORES	NH	L	R	FC	C	SH	M	I	H	A	AT	SX	AB	AX	MS	BR	PN	P
	45	56	5	96	1	1	21	3	25	14	2	0	2	7	6	5	2	9

Examples of computer input and output

APPENDIX F

COVER LETTER TO STAFF

MICHIGAN STATE UNIVERSITY EAST LANSING • MICHIGAN 48823

COLLEGE OF EDUCATION • OFFICE OF STUDENT TEACHING • ERICKSON HALL

February 27, 1968

Name
Address
Address
Address
Address

Dear ,

Thanks for your help in gathering the Holtzman Inkblot Records in your center. In all, I collected 815 records and received summary scores on 18 perceptual modes for each student teacher.

The next phase of the study is to gather success ratings of all the student teachers to determine a relationship between how they see things and judgments of their success.

Enclosed are SUCCESS RATING CHARTS which explain the meaning of each success rating number. Attached to each CHART is a SUCCESS RATING FORM which you can use to list the student's name and his success rating number. I hope I have enclosed enough charts and forms but more are available if you need them.

The study, of course, is only as good as the careful rating of each student. And the most crucial question the study considers is the perceptual pattern of successful student teachers.

Term end time is extremely busy in all the centers and I am grateful for your help.

I am enclosing a return self-addressed stamped envelope for your rating sheet (no need to return charts). The return envelope is addressed to my home since mail usually comes there sooner than at the office.

Once again, thanks very much and I hope to report the findings at Gull Lake on April 24th.

Sincerely,

B. Bradley West

APPENDIX G

SUCCESS RATING FORM

SUCCESS RATING FORM

Rater's Name _____ Center _____ Date _____

STUDENT'S NAME	SUCCESS RATING NUMBER	STUDENT'S NAME	SUCCESS RATING NUMBER