Characteristic Differences, Determined By Tat Sequential Analysis, between Teachers Rated By their Pupils At the Extremes in Teaching Efficiency

Mary Innocentia Burkard
Loyola University Chicago
CHARACTERISTIC DIFFERENCES, DETERMINED BY TAT SEQUENTIAL ANALYSIS, BETWEEN TEACHERS RATED BY THEIR PUPILS AT THE EXTREMES IN TEACHING EFFICIENCY

by

Sister Mary Innocentia Burkard, SSND

A Dissertation Submitted to the Faculty of the Graduate School of Loyola University in Partial Fulfillment of the Requirements for the Degree of Doctor of Philosophy

June 1958
Sister Mary Innocentia Burkard was born January 3, 1909, in Green Bay, Wisconsin, where she completed her elementary and secondary education at St. Francis Xavier Cathedral Grade and High School and St. Joseph Academy.

In 1927 Sister entered the Community of the School Sisters of Notre Dame in Milwaukee. She received the Bachelor of Arts degree from Mount Mary College there and, in 1949, the Master of Arts degree from Marquette University. Prior to her coming to Loyola in 1955, Sister also attended summer sessions at Notre Dame, St. Louis, and Creighton Universities, and the Catholic University of America.

Sister Mary Innocentia taught at St. Anthony School, Milwaukee; Visitation School, Elm Grove, where she was also principal; and at Messmer High School, Milwaukee. From 1949 to 1955 she was Community and Diocesan School Supervisor in the Diocese of La Crosse, Wisconsin.

Sister wrote her Master's thesis on "Teacher Participation in Catholic School Administration," parts of which were published under the same title in the Catholic School Journal, XLIX (October and November, 1949), 261-263; 299-301. She is also the author of "The Catholic Teacher's Responsibility for Reading Guidance," Catholic Library World, XIX (February 1948), 147-149.
ACKNOWLEDGMENTS

The author wishes to express her gratitude to the members of Loyola's Education and Psychology Departments for the encouraging help they so consistently extended; particularly to Dr. Magda B. Arnold for her patient guidance in TAT analysis. Very special thanks are due to the generous School Sisters of Notre Dame, over three hundred strong, who so graciously allowed themselves to be rated and tested in the interests of Virtus et Scientia.
TABLE OF CONTENTS

LIST OF TABLES AND FIGURES ............................................................ v

Chapter

I. THE PROBLEM .................................................................................. 1

Brief statement of the problem--Background of the problem--Attempts to find differentiating characteristics of teachers in intelligence, observed behavior, inventories, and projective techniques--Rationale of the present study--Plan of the present study.

II. INVESTIGATIONAL MEANS AND PROCEDURES ............................ 23

Description of the instruments--The subjects of the investigation--Procedures in collecting the data.

III. ANALYSIS OF THE DATA ............................................................... 37

Analysis of the ratings--Investigation of the whole sample of three hundred teachers: study of the Otis Test and of the MTAI--Investigation of the paired sample: bases of pairing--Study of the MTAI and of the TAT.

IV. INTERPRETATION OF THE TAT FINDINGS ................................. 67

Detailed description of the findings--The nature of TAT findings--Teacher characteristics as evaluated by the pupils.

V. DISCUSSION .................................................................................... 84

General conclusions--Limitations and needed research--Implications for teacher education.

VI. SUMMARY ...................................................................................... 89

Specific results--Values.

BIBLIOGRAPHY .................................................................................. 94

APPENDIX ............................................................................................ 102
# LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>I. AGE, TEACHING EXPERIENCE, AND TEACHING LEVELS OF TEACHERS</td>
<td>31</td>
</tr>
<tr>
<td>PARTICIPATING IN THE STUDY</td>
<td></td>
</tr>
<tr>
<td>II. NUMBER OF PUPILS WHO RATED THEIR TEACHERS ON EACH LEVEL</td>
<td>31</td>
</tr>
<tr>
<td>III. DISTRIBUTION OF TEACHERS AND PUPILS BY GRADE GROUPS</td>
<td>32</td>
</tr>
<tr>
<td>IV. INTERCORRELATIONS OF THE SEVEN RATINGS (High School)</td>
<td>39</td>
</tr>
<tr>
<td>V. INTERCORRELATIONS OF THE SEVEN RATINGS (Grade School)</td>
<td>40</td>
</tr>
<tr>
<td>VI. MEASURES OF CENTRAL TENDENCY AND VARIABILITY OF INTELLIGENCE</td>
<td>42</td>
</tr>
<tr>
<td>TEST SCORES OF ENTIRE SAMPLE</td>
<td></td>
</tr>
<tr>
<td>VII. RELATIONSHIP BETWEEN INTELLIGENCE SCORES OF HIGH SCHOOL teachers and their rating on each of seven questions</td>
<td>43</td>
</tr>
<tr>
<td>VIII. DIFFERENCES BETWEEN MEANS ON THE MTAI OF CONTRASTING CRITERION</td>
<td>45</td>
</tr>
<tr>
<td>GROUPS ON VARIOUS GRADE LEVELS</td>
<td></td>
</tr>
<tr>
<td>IX. DIFFERENCES BETWEEN HIGH AND LOW GROUPS ON CATEGORIES OF THE</td>
<td>48</td>
</tr>
<tr>
<td>MTAI (Grade 4)</td>
<td></td>
</tr>
<tr>
<td>X. DIFFERENCES BETWEEN HIGH AND LOW GROUPS ON CATEGORIES OF THE</td>
<td>48</td>
</tr>
<tr>
<td>MTAI (Grades 5-6)</td>
<td></td>
</tr>
<tr>
<td>XI. DIFFERENCES BETWEEN HIGH AND LOW GROUPS ON CATEGORIES OF THE</td>
<td>49</td>
</tr>
<tr>
<td>MTAI (Grades 7-8)</td>
<td></td>
</tr>
<tr>
<td>XII. DIFFERENCES BETWEEN HIGH AND LOW GROUPS ON CATEGORIES OF THE</td>
<td>49</td>
</tr>
<tr>
<td>MTAI (Grades 9-10)</td>
<td></td>
</tr>
<tr>
<td>XIII. DIFFERENCES BETWEEN HIGH AND LOW GROUPS ON CATEGORIES OF THE</td>
<td>50</td>
</tr>
<tr>
<td>MTAI (Grades 11-12)</td>
<td></td>
</tr>
<tr>
<td>XIV. RELATIONSHIP OF AGE TO INTELLIGENCE SCORES, MTAI SCORES, AND</td>
<td>52</td>
</tr>
<tr>
<td>RATING (Grades 4-8)</td>
<td></td>
</tr>
</tbody>
</table>
Table

XV. RELATIONSHIP OF AGE TO INTELLIGENCE SCORES, MTAI SCORES, AND RATING (Grades 9-12) ... 52

XVI. COMPARATIVE STATISTICS OF THE PAIRED GROUPS ON AGE ... 54

XVII. COMPARATIVE STATISTICS OF THE PAIRED GROUPS ON INTELLIGENCE. ... 55

XVIII. COMPARATIVE STATISTICS OF THE PAIRED GROUPS ON RATING: ... 55

XIX. DIFFERENCES ON THE MTAI BETWEEN PAIRED SECONDARY GROUPS ... 57

XX. DIFFERENCES ON THE MTAI BETWEEN PAIRED ELEMENTARY GROUPS ... 57

XXI. DIFFERENCES BETWEEN PAIRED ELEMENTARY GROUPS ON EIGHT ITEMS OF THE MTAI ... 59

XXII. ILLUSTRATIVE STATEMENTS FROM TWO CATEGORIES OF THE SCORING STANDARD ... 62

XXIII. SCORING RELIABILITY OF THE TAT ... 63

XXIV. SCORES OF THE HIGH AND LOW GROUPS ON THE TAT ... 64

XXV. RELATIONSHIPS BETWEEN TAT SCORES AND RATINGS ON THE SEVEN QUESTIONS OF THE SCALE ... 65

XXVI. HABITUAL BASIC DISPOSITION AS EXPRESSED TOWARD SUCCESS ... 69

XXVII. HABITUAL BASIC DISPOSITION AS EXPRESSED TOWARD FAILURE ... 70

XXVIII. HABITUAL BASIC DISPOSITION AS EXPRESSED TOWARD LOSS ... 71

XXIX. HABITUAL BASIC DISPOSITION AS EXPRESSED TOWARD LIFE AND ITS OBLIGATIONS ... 72

XXX. HABITUAL BASIC DISPOSITION AS EXPRESSED TOWARD OTHER PEOPLE. 73

Figure

1. CENTILE EQUIVALENTS OF AN AVERAGE RATING OF 4.0 FOR TEACHERS ON VARIOUS GRADE LEVELS ... 38
CHAPTER I

THE PROBLEM

This study is an attempt to determine differences between teachers who are ranked by their pupils at the extremes in teaching efficiency. It is a search for fundamental personality characteristics that will clearly discriminate, it is hoped, between contrasting criterion groups, and that will be intelligible on the basis of a logical connection between the personality of a teacher and her effect on other personalities.

If such differences can be found between those who make the most favorable impression on their pupils and those who make the most unfavorable impression, the findings should be of value for further research in teaching efficiency and should point the way to practical conclusions regarding most fruitful procedures in the guidance and training of student teachers.

Background of the Problem

During the past fifty years hundreds of studies have wrestled with the problem of teacher efficiency and inefficiency from various points of view, and today interest in the subject is not abating. Watters' recent annotated bibliography describes ninety-nine reports of such studies between the years 1949 and 1954;¹ and the latest summary in the Review of

Educational Research lists eighty-six between June 1952 and June 1955.2

Most of these studies have contributed to our knowledge of teachers' qualities, both good and bad; or they have discovered what kind of teachers are wanted by administrators, supervisors, and pupils; or they have developed new techniques for measuring factors considered to be evidences of, or contributing to, success in teaching. A vast quantity of material has accumulated, some of it supporting, some of it contradicting previous research. And yet, on top of this accumulation of data appear such bold, discouraging statements as, "Few, if any, statistically reliable findings have been reported that help to answer the question, 'What are effective teachers like,'"3 and, "It is a peculiar circumstance that, despite the critical importance of the problem and a half-century of prodigious research effort, very little is known for certain about the nature, measurement, and prediction of teacher effectiveness."4

What are the reasons for these disappointing conclusions after the expenditure of so much effort?

Any study that attempts to analyze the characteristics of an effective teacher must begin by deciding two things:

(1) Where are we to look for the distinguishing characteristics of the

---


3David G. Ryans, "The Investigation of Teacher Characteristics," The Educational Record, XXXIV (October 1953), 379.

effective teacher? Is it in her preparation by way of scholarship, educational achievement, or professional training? Is it in her outward bearing or deportment; in the things she does or does not do in the classroom? Is it in this external, observable area that the differences are to be found between the effective and the ineffective teacher? And if so, is an effective teacher made, therefore, by concentrating on the development of these exterior practices? Or is there another level of personality on which the determining factors of the effective teacher are to be found?

(2) How are we to say when a teacher is effective? Can we tell by observing the teacher in action? Can the principal tell us, or the supervisor? Can we find out by examining the pupils? Can the pupils themselves tell us? Or do we perhaps have to wait until the pupils have grown up and display in their living the accumulated learning derived from the action and interaction of and with their many teachers both in and out of school? Obviously, to accept the last as the only legitimate criterion would be to eliminate the possibility of ever discovering the relative effectiveness of any particular teacher.

Researchers studying teacher efficiency have made these decisions in different ways. Since the present study is expected to contribute primarily to the area represented by the first set of questions, it is that area which will be our chief concern here.

Broadly speaking, there are four sources to which research has looked for the distinguishing characteristics of good and poor teachers. These sources are: the teacher's intelligence and educational preparation; the teacher's observed behavior in the classroom; the attitudes and other
personality characteristics expressed by the teacher in inventory and questionnaire responses; and finally, underlying basic personality elements revealed in disguised instruments of evaluation. According to their theoretical assumptions, researchers have sought in one of these areas, or in a combination of two or more of them, those qualities that would clearly differentiate the good from the bad as defined by their accepted criterion of efficiency.

Differences in Intelligence or Preparation

A comparatively large group of researchers in teaching efficiency have included among their variables a measure of intelligence, with somewhat inconsistent results. Rostker, for instance, concluded that, of all the factors he explored, intelligence seemed to have the highest relationship to successful teaching.5 Likewise, La Duke and Bendig found significant correlations between intelligence and competency as measured by their criteria.6 On the contrary, a number of other studies, notably those of Rolfe, Bach, and Dodge, show relatively low correlations between teaching success and intelligence of elementary and secondary teachers.7


7J. F. Rolfe, "The Measurement of Teaching Ability," Journal of Experimental Education, XIV (September 1945, 65; Jacob O. Bach, "Practice Teaching Success in Relation to Other Measures of Teaching Ability,"
Though the evidence is somewhat in favor of a relationship between successful teaching and intelligence, the relationship is not sufficiently high to warrant a discrimination between efficient and inefficient teachers on the basis of intelligence alone. The question arises, of course, as to whether instruments designed to measure scholastic aptitude are adequate evaluations of a teacher's intelligence.

Neither has educational nor professional preparation been found to distinguish between good and poor teachers. The most favorable conclusion, reached, for instance, by Lins, Rostker, and Drawhorne, is that scholastic achievement, as measured, is important or significant to some degree. On the other hand, Seagoe and Carlile found very low correlations between ratings and achievement test scores of elementary and secondary student-teachers. La Duke and Bach found professional preparation not significant on the elementary and the secondary levels respectively; and Ryans concluded that amount of college training appeared to be but slightly

---

Journal of Experimental Education, XXI (September 1952), 75-78; Galen W. Dodge and Donald O. Clifton, "Teacher-Pupil Rapport and Student-Teacher Characteristics," Journal of Educational Psychology, XLVII (October 1956), 365-370.


La Duke, p. 94; Bach, p. 78.
related to teacher performance, and that there was even some evidence in favor of smaller amounts of such training: 11

Thus, we may safely say that the distinguishing characteristics of efficient and inefficient teachers have not been found in the measured intelligence and knowledge of the teachers, nor in the quantity of preparation that they have received.

Differences in Observed Behavior of Teachers

A part of the research on teacher efficiency seeks to find the distinguishing qualities of good teachers in their behavior in the classroom as observed by administrators, supervisors, education professors, and other experts. Such, for example, is the 1929 Barr study, which resulted in a list of characteristics of teachers in the form of practices and traits observed. 12 The Commonwealth Study of the same year presents a similar list extracted from common classroom practices and based on the opinions of selected judges. 13 More recently Jensen, in a part of the Teacher Characteristics Study, collected "critical behaviors", that is, samples of observed behavior which appeared to be peculiar to either the effective or the ineffective teacher. 14

11 Ryans, p. 391.


While these studies have resulted in the listing of desirable characteristics of teachers such as fairness, cheerfulness, sympathy, and skill, they fail to discriminate between good and poor teachers. Methodologically speaking, many of the studies in this group suffer from the defect of having the same data, that is, observed external conduct, serve as both independent and dependent variables.

However, those studies that have used a different criterion have been no more successful. Ratings by pupils have been employed, sometimes alone and frequently in conjunction with adult ratings. One of the better recent studies using pupil ratings as criteria is that of Symonds whose findings point to more basic characteristics such as personal organization, feelings of security, and apparent integration of personality. These are, however, inferences made subjectively from observation of external practices, a fact which is recognized by the author, who looks upon this as an exploratory study preparing the way for more definitive research. 15

Some of the researchers who have sought the distinguishing characteristics of teachers in surface qualities have attempted to improve their studies by using as criterion the effect of teaching on the pupils, as indicated by specific measurable outcomes, or observed reaction of pupils and interaction between teacher and pupils. Here again results are inconclusive. Lancelt found a combination of desirable personality traits and methods in the teachers who rated highest in the specific outcomes measured,

namely, subject grades in subsequent courses.\textsuperscript{16} Anderson and Baxter likewise observed the more desirable traits in the teachers who had the most beneficial effect on the attitudes, reactions, and personalities of pupils.\textsuperscript{17} All of these studies dealt with very small samples (the largest was thirteen teachers); and they did not account for those teachers who demonstrated desirable qualities when observed and still ranked low on the criterion measure, nor for those who did not demonstrate desirable qualities and yet ranked high.

In their somewhat larger study, La Duke and Rostker found little relationship between their criterion of pupil change and the qualities and practices of the teachers as rated by supervisors.\textsuperscript{18} More recently McCall, using a composite weighted measure of pupil growth over a period of time, reported slightly negative correlations with ratings by principals based on observations.\textsuperscript{19}

On the whole, one is forced to conclude that, regardless of the criterion employed, the distinguishing characteristics of efficient and inefficient teachers have not been found by the use of behavior sampling techniques.


\textsuperscript{17}Harold H. Anderson and S. E. Brewer, \textit{Studies of Teachers' Personalities}, II (Berkeley, Cal., 1946), pp. 124-125; Bernice Baxter, \textit{Teacher-Pupil Relationships} (New York, 1943), pp 32-117.

\textsuperscript{18}La Duke, p. 99; Rostker, p. 50.

\textsuperscript{19}W. A. McCall, \textit{Measurement of Teacher Merit} (Raleigh, N. C., 1952), cited by Watters, p. 361.
Differences in Inventory and Questionnaire Responses

It would naturally be expected that attitude and trait inventories would penetrate somewhat more deeply into underlying aspects of the teacher's personality than observation of external behavior and thus produce more discriminating results. Numerous instruments have been used, and even specifically created, to find the fundamental differences between good and poor teachers. In reviewing these studies it is important to distinguish between those that have employed such instruments merely for purposes of prediction and those that have been interested in discovering traits.

The former studies are not so much interested in what they measure as in what they predict; that is, attention is directed toward certain specific responses which have been found by empirical tests to predict some specific criterion behavior. Such studies are of little value in defining the psychological processes or functions of efficient teachers. In this group is, for example, the research with the Minnesota Multiphasic Personality Inventory, such as that of Gough and Pemberton, Michaelis, Tanner, and Tyler.20 Even if these studies had succeeded in discriminating between good and poor teachers, they would have told us nothing about the actual characteristics of these teachers.

The second type of studies uses instruments that have been built on

rational conceptions of traits or that have been developed from more or less homogeneous clusters of behaviors that are presumed to constitute traits. These include the older inventories, such as the Bernreuter and the Bell, and the newer batteries developed through factor analysis, such as the Guilford-Martin, the Thurstone Temperament Schedule, and the Cattell Questionnaires.

Results of studies with any of these instruments have been disappointing as far as discrimination is concerned. Jones found a correlation of -.04 between supervisors' ratings and scores on the Bell Adjustment Inventory.\(^{21}\) Laycock's coefficients ranged from -.21 to +.33 between ratings of success and measures from the Bernreuter Personality Inventory.\(^{22}\) In a more recent study Carlile found neither of these instruments yielding notable results.\(^{23}\) Gotham found no significant relationships between his criterion of pupil change and scores on the Bernreuter, the Washburne Social Adjustment Inventory, and the Rudisill Scale for the Measurement of the Personality of Elementary School Teachers.\(^{24}\) While Tanner's results show some correlation between student-teachers' ratings and the YWCA Secretary Scale of the Strong Vocational Interest Blank and the Social Service


\(^{22}\) S. R. Laycock, "The Bernreuter Personality Inventory in the Selection of Teachers," *Educational Administration and Supervision*, XX (January 1934), 59-63.

\(^{23}\) Carlile, p. 667.

Scales of the Kuder Preference Record, he stated that there was much overlapping of traits and no discrimination between good and poor teachers.\textsuperscript{25}

Ryans discovered nothing more than a tendency for three scales of the Thurstone Temperament Schedule to distinguish between criterion groups, namely, the dominant, the sociable, and the impulsive;\textsuperscript{26} while Bendig found a similar trend with the Guilford-Zimmerman Temperament Survey.\textsuperscript{27} Two studies based on the Primary Source Traits developed by Cattell proved no more successful. Schwartz reported some relationship with reaction time and two-hand coordination;\textsuperscript{28} and Lamke concluded that good teachers are good for different reasons and poor teachers fail for varying reasons, but no adequate patterns for such success or failure were forthcoming in his study.\textsuperscript{29} Still more recently Erickson and Montross, working independently with the Thurstone and Cattell Scales, failed to find significant relationships with teaching success as measured by a variety of criteria.\textsuperscript{30}

Another group of studies dealing with teachers' attitudes toward

\textsuperscript{25}Tanner, p. 274.

\textsuperscript{26}Ryans, p. 391.

\textsuperscript{27}Bendig, p. 706.


\textsuperscript{29}T. A. Lamke, "Personality and Teaching Success," \textit{Journal of Experimental Education}, XX (December 1951), 253-254.

factors directly connected with school and teaching were somewhat more conclusive in their findings. Cook, Leeds, and Callis found significant relationships between scores on the Minnesota Teacher Attitude Inventory (MTAI) and combined ratings of experts, principals, and pupils. Coefficients here were as high as .63 between Inventory scores and the three criteria with combined multiple weights.\(^31\) Kearney and Rocchio reported the same instrument useful for predicting feeling tone on the part of the pupil for his teacher.\(^32\) In a part of the Teacher Characteristics Study, Wandt concluded that teachers rated high and low by their principals differed significantly in their attitudes toward pupils and toward administrators.\(^33\)

It would appear that these measures of teachers' attitudes toward their immediate contacts in their teaching task have more promise with regard to revealing at least some of the characteristics of the successful teacher. Their adequacy for this purpose will depend on the degree to which they can be made more reliable and the dangers of faking can be reduced. Studies by Eson and by the authors of the MTAI have shown that the responses of the teachers to inventories are not necessarily a true expression of


their attitudes. Recognizing this, the workers in the Teacher Characteristics Study have, to a great extent, directed their research toward finding and creating disguised instruments with which to arrive at true dimensions of teacher personality. Up to the present, however, reports on the use of the Teacher Characteristics Schedule, containing verbal items and pictorial materials, have not differed notably from previous research.

**Differences in Responses to Projective Techniques**

Attempts to seek the differentiating characteristics of teachers in deeper levels of the personality by means of projective techniques have been increasing. Guba and Getzels reported a comparative study in which they used two normative instruments, the Guilford-Martin Inventory and the Allport-Vernon-Lindzey Study of Values, and two projective devices, the Rosenzweig Picture-Frustration Study and a sentence completion test prepared by the authors. More significant relationships were found with the projective instruments than with the normative. The results are, however, clouded by the ambiguity of the terminology and the subjective conceptions with which they are interpreted.

Alexander, using a specially prepared form of the Thematic Apperception

---


Test that features children and adults, explored the possibility of predicting behavior patterns of teachers in interaction with children. He reported that 76 per cent of his predictions coincided with the observational data and that reliability was high, a second analyst agreeing with 89 per cent of his ratings. While this study was not designed to discriminate between criterion groups, it represents the highest degree of accuracy in prediction that has been reported, and thus it indicates the possibilities inherent in projective tests, particularly the TAT.

Two other studies are reported using the Alexander version of the TAT. Ohlsen and Schulz found several significant differences between high and low-rated student-teachers when their stories were analyzed on the basis of eight questions prepared for the purpose. Oelke, using both the MTAI and the Alexander TAT, found that the TAT discriminated in amount and direction of change between small numbers of best and poorest student-teachers during their initial teaching experiences, whereas the MTAI did not. However, he concluded that results from neither were sufficiently significant to warrant their use as predictive instruments at this time.

Johnson likewise used a projective test of the TAT type composed of

---


ten pictures depicting situations assumed relevant to educational adjustment. The relationship he found between the TAT scores and his criterion of specially designed observation ($r = .75$) led him to conclude that the TAT appeared to be a valuable tool for predicting significant aspects of teaching effectiveness. $^{40}$ His results with the Rorschach in the same study were less conclusive. Similarly, Cooper, Page, and Travers discovered no relationships between quantitative Rorschach factors or triads of Rorschach ratios and the ratings of teachers by supervisors or pupils. $^{41}$

The above researches with the TAT reported reliabilities of scoring that ranged from 80 percent (Ohlsen and Oelke) through 89 percent (Alexander) to 91 percent (Johnson). Alexander used both analysis of form characteristics and content of stories to study seven categories, all but one of which he found significant at the one or two percent level. Ohlsen and Oelke likewise used a content analysis. Johnson's scoring was done on the basis of a five-point scale, on which credit was given from the standpoint of ability to find the problem suggested by the picture and the ability to solve the problem once found. It is highly probable that a great deal of subjectivity entered into the analyses in all of these TAT studies which prevented them from producing more conclusive results.


Rationale of the Present Study

The above brief overview of the literature reveals that there has been no dearth of attempts to answer the question, "How do good teachers differ from poor ones?" It also demonstrates that the answer has not been found. Research workers have attributed this to the lack of an adequate criterion of teaching efficiency. It has been said, for instance, that the validity of our assumptions and opinions regarding teaching cannot readily be tested because there is so little understanding, and there are no adequate measures, of the criteria of teacher effectiveness.42

In line with this thinking is the conclusion, appearing repeatedly in the literature on teacher effectiveness, that competency is not a global thing, but is made up of specific, objectively-defined behaviors that are relative to the situation in which they appear. Thus the search for a criterion today is concerned not with teaching effectiveness as a whole, but with such questions as: what kind of teachers achieve what kind of goals with what kind of children in what kind of situations.43

The position taken here is that it is not necessary to await a more adequate criterion of teaching efficiency before proceeding with research


in the characteristics of good teachers. In every profession or art it is possible to determine who is successful or unsuccessful in certain aspects at least. The same is true of the teacher. He may be considered efficient by the administrator or supervisor; he may be rated high by his pupils; or he may be successful in bringing about changes of various kinds in the pupils.

Not all of these criteria may be equally discriminating nor equally reliable. But whatever criterion of success is used, there are likely to be real differences between those who are high and those who are low on the criterion measure, provided the latter is sufficiently discriminating. Though the contrasting data may vary with different criteria, it should be possible to arrive at real distinctions among them. What is needed, however, is an adequate instrument with which to evaluate the personalities in the contrasting criterion groups.

Even though a variety of apparently significant criteria have been used in the reported studies, no differentiating characteristics have been found. This is true in spite of the fact that researchers have taken care to make their criteria trustworthy. In some studies there was strong agreement among administrators and supervisors as to the success of the teachers rated. There was even more consistency among the thousands of pupils who rated their teachers. And still no really distinguishing qualities appeared in these obviously different criterion groups. In spite of these failures it is safe to say that the differences are there; and that if they have not been identified, it may be due to the sources in which, or the means by which, the differences have been sought.
In the present study, the adequacy of a previously tested criterion, that is, rating by pupils, is accepted; but a new departure is made in the selection of the predictive instrument. It is suggested that the differences between successful and unsuccessful teachers do not lie in external, observable behavior. Nor are they to be found in the areas probed by the ordinary personality inventory. The differences lie in more fundamental elements of the teacher's personality; and these elements can be reached only by a disguised technique which is objectively interpreted, so that the prejudices and prior assumptions of neither the testee nor the tester will intrude to distort the results.

Plan of the Present Study

The Criterion

The criterion selected is the one which appears to this investigator the most satisfactory both as to validity and as to reliability. Numerous studies have attested to the fact that pupils can express opinions about the efficiency of teaching, and that they do so in a way that is both forthright and discriminating.44

There is no intention here of demonstrating the validity of pupil rating as a measure of complete teacher success, whatever that may be. The position taken is that the rating received by a teacher from her pupils, if it is procured in such a way that it can be considered the true opinion of the raters, is in itself an indication of success or failure to achieve something—in this case, the esteem of her pupils; and the groups discriminated by such rating are likely to have different characteristics. Such pupil rating must, then, for the purposes of this study, be accepted as valid.

As to reliability, there is adequate evidence in the research literature that, in their ratings, pupils are consistent with one another and with themselves. In his study with 1500 pupils of grades eight to eleven, Bryan found pupil ratings highly reliable ($r = .90$ and $.95$), and their self-consistency practically perfect ($r = .995$ and $.997$). Hart concluded from his survey of 10,000 high school students that pupils are mature enough to weigh values and arrive at reliable and significant evaluations of their teachers. Cook and Leeds found reliability coefficients of $.93$ in their study with middle grade pupils. Beecher reviewed studies extending over a period of twenty years and involving a total of more than thirty thousand boys and girls in elementary and secondary schools. 

---


emphasized that the consistency of pupils with each other, even of pupils of widely varying social, economic, and racial backgrounds, was much greater than that of adult raters. He stated that the thirteen thousand pupils whose voluntary reactions to their teachers were obtained by Davis in his 1924 study might as well have been the white and colored groups of the 1943 investigation carried on in South Carolina, as far as their evaluation of their teachers' characteristics was concerned. 48

In view of all these findings, given a trustworthy method of securing the ratings, we can reasonably accept them as both valid and reliable.

The Predictive Instruments

Since previous research has given some indication of a significant, though not discriminating, relationship between success in teaching and intelligence, and between success in teaching and expressed attitudes of the teacher toward pupils and other factors in the teaching situation, it was considered advisable to further explore these two areas, both for their own sake and for what they might contribute by way of controls to the remainder of the investigation. Accordingly, two instruments were selected to measure these factors, the Otis Quick-Scoring Mental Ability Test and the Minnesota Teacher Attitude Inventory (MTAI).

The most significant differences between the teachers were expected to be found, however, as was stated above, by a disguised technique that could be objectively scored. For this purpose the investigator selected

the Thematic Apperception Test (TAT) as most promising for its penetration, and TAT Sequential Analysis as the most objective way of analyzing and scoring. Being a disguised type of personality test, the TAT is not as easily faked as are the existing kinds of attitude and personality inventories. Of all projective techniques, it appeared most suitable to the present research because it can indicate the deeper motivational levels of the personality. TAT Sequential Analysis, based as it is on the objective reproduction of the significance of the stories told, is free from the subjective interpretation and the ambiguities of some other methods of TAT analysis.

The Subjects

To reduce the number of intervening variables, it was decided to restrict the sample to teachers belonging to one community of Catholic Sisters, who are fairly homogeneous because they have a common training and environment. Since past research has raised the question as to whether the teacher who is rated high by young pupils differs characteristically from the one who is rated high by older students, this question was also made a part of the present investigation. Hence, an equal number of elementary and of secondary teachers was included.

Purposes

It is hoped that the results of the study will be useful to further research in this field. It was also hoped that the findings would be of such a nature as to contribute to the improvement of guidance, training, and placement programs for student teachers.
Specific Problems

On the basis of the above rationale, the specific problems to be explored, then, may be succinctly put as follows:

(1) Are there statistically significant and discriminating differences with respect to intelligence, as measured by the Otis Quick-Scoring Mental Ability Test, and with respect to professional attitudes, as indicated by the Minnesota Teacher Attitude Inventory, between:
   (a) the high-rated and the low-rated elementary teachers;
   (b) the high-rated and the low-rated secondary teachers;
   (c) the elementary and the secondary teachers of either rating?

(2) Are there statistically significant and discriminating differences between any of the above criterion groups on data derived from the Thematic Apperception Test?

(3) If there are significant differences on the scales or in the TAT material, how can these differences be interpreted and used in the guidance, training, and placement of student teachers?

---

49 Throughout this study the term "discriminating differences" is used to denote clearly differentiating characteristics of criterion groups, in a practical, rather than a merely statistically significant, sense.
CHAPTER II

INVESTIGATIONAL MEANS AND PROCEDURES

The thinking that directed the choice of instruments for this study has been outlined. The present chapter will deal with methods used in collecting the data and will describe in greater detail the tools and techniques employed.

Description of the Instruments

The Rating Scale

The rating instrument was the Diagnostic Teacher-Rating Scale, Form A, developed by Sister Mary Amatora. The scale consists of a short form called the Area Scale including the seven divisions: (1) liking for teacher; (2) teacher's ability to explain; (3) kindness, friendliness, and understanding; (4) fairness in grading; (5) discipline; (6) amount of work required; (7) liking for lessons. Pupils are asked to rate the teacher on each area according to a five-point scale. Following this is a Diagnostic Check List consisting of forty-nine scaled statements. The items are arranged in seven intra-scales of seven items each. These items are in turn scaled so that their values extend from highest to lowest, with the fourth

---

1Sister Mary Amatora, "A Diagnostic Teacher-Rating Scale," *Journal of Psychology*, XXX (October 1950), 396-399.
item in each intra-scale at the midpoint. Pupils mark the statements with which they agree.

Reported reliabilities, computed by the split-half method, for the seven area scales range from .86 to .96. For the diagnostic intra-scales, Form A vs. Form B, they range from .87 to .97 when corrected for attenuation according to the Spearman-Brown formula. Validity is argued by the logic underlying the scale construction—that truthful, reliable expressions of opinion are valid.2

The scale was originally developed with elementary school pupils. To determine whether it would be suitable for the secondary school pupils included in the present research, a preliminary study was made with a group of 146 high school pupils. The latter were asked to rank, according to their estimate of importance, thirty-four characteristics of teachers, half of which were taken from the Diagnostic Teacher-Rating Scale and the other half from various other rating scales which had been prepared specifically for high school pupils. Of the seventeen items taken from Sister M. Amatore's scale, thirteen were ranked by the high school students in the upper half of the list, and all of the broad items of the Area Scale were ranked high. The scale was, therefore, considered adequate for the secondary school pupils as well as for the elementary, and was used for both groups, thus facilitating comparisons.

2Ibid., 397.
The Intelligence Test

The Otis Quick-Scoring Mental Ability Test, Gamma, Form Am, was considered most suitable for the present investigation. The fact that the Otis Test has but a single time limit made it less threatening to the older adult subjects, many of whom had never before taken an intelligence test. While a longer test might have had advantages, the length of the testing session had to be kept within reasonable bounds, and the half-hour time limit of the Otis was considered most desirable.

The authors report reliability coefficients ranging from .85 to .91 for Form Am. Evidence for validity is based on item-selection procedures; coefficients have a median value of .61. Studies have indicated a high relationship between the Otis tests and the American Council on Education Psychological Examination, with the former showing as high, or higher, relationship to college grades. Many validation studies in which the Otis test was checked against industrial and professional criteria have yielded significant validity coefficients. While it is recognized that this test does not discriminate clearly at the upper levels, the discrimination was considered adequate for the purposes of this study.

The Teacher Attitude Test

The Minnesota Teacher Attitude Inventory (MTAI) developed by Cook, Leeds, and Callis, is designed to measure those attitudes of a teacher which predict how well he will get along with pupils in interpersonal relationships and was, therefore, considered particularly applicable to the present investigation. It consists of 150 items which have been found by the authors to discriminate sharply between teachers who have, and those who do not have, good rapport with pupils. The authors report validity coefficients of .49 when pupil ratings were used and .63 on a composite criterion of principals', experts', and pupils' ratings. Reliability coefficients of .93 have been consistently produced.6

The possible range of scores on the MTAI is from plus 150 to minus 150, the "right" and "wrong" answers having been determined empirically with groups of teachers on the basis of their teacher-pupil relations. It was recognized that the scores on the MTAI reflect the educational philosophy of the authors and of the validation groups and that their philosophy may differ considerably from that of the subjects in the present study. The investigator, however, in a preliminary try-out of the Inventory, made a score that ranked at the 99th centile of the highest norm group of experienced teachers. It was concluded, therefore, that differences in philosophy would not invalidate the inventory for the present group, to which the investigator herself belongs.

The Projective Test

The Thematic Apperception Test (TAT), as originated by Murray and Morgan, consists of a series of pictures about which the subject is asked to tell stories. There are four overlapping sets of twenty cards for boys, girls, men, and women. In the present investigation eleven pictures and the blank card were selected for presentation in the following order: 1, 2, 3BM, 4, 5, 6BM, 7BM, 8EM, 9BM, 10, 16, 11.

A dozen or more systems of analyzing and scoring the test have developed in addition to Murray's original one.\(^7\) Of these the method of TAT Sequential Analysis was chosen as the most objective and the most theoretically sound. Originated by Dr. Magda B. Arnold,\(^8\) this method consists essentially in abstracting from the story its full import as revealed by the plot and its outcome. Every story, as told by its author, expresses a certain orientation, a way of looking at life, self, or others. The author, taken up with the details of his story, is not fully aware of this philosophy to which he is giving expression and which is actually a strong motivating power in his life. The psychologist, however, upon reading each story can penetrate to its meaning and can set down in a succinct statement what the writer expresses through his story plot and its outcome. The result of this analysis is not a subjective interpretation on the part of the psychologist, but merely a restatement in a generalized, abstract form

---

\(^7\) Edwin S. Shneidman et al., *Thematic Test Analysis* (New York, 1951).

of what the writer is saying by means of a particular, concrete situation.

It is characteristic of these generalized statements always to follow a sequence. If the abstraction is correctly done, there will appear an association among these statements from story to story, and this feature makes of the total protocol a more or less continuous, connected expression of the subject's way of looking at his world and of handling his problems. This sequential feature of the TAT protocol has been discovered empirically in hundreds of TAT analyses, and the fact of its existence has been repeatedly affirmed in subsequent discussions with the subjects concerned. While other experts in TAT analysis have likewise found a tendency in TAT stories to be related,9 this method of abstracting the import of the story reveals such a sequence of ideas to be an unfailing phenomenon underlying every series of TAT stories. This does not mean that one theme is necessarily carried throughout the series. Depending upon the number of pictures used, there may be two or more themes formed by clusters of stories and usually more or less related.

The existence of this sequence serves as a guide in the TAT analysis. It happens at times that a subject may appear to be saying several things through his story. While all he says may be true expressions of his philosophy, the one that best fits the sequence will be the most relevant to his mental set and emotional disposition at the time of writing the stories. This fact has also been repeatedly demonstrated by clinical work with the subject following a TAT analysis.

The value of this method for research purposes lies particularly in its objectivity, which is high in comparison with other systems. To the degree that the ideas, the attitudes, and the theoretical assumptions of the analyst are injected into a test, the picture of a given personality is distorted. Objectivity requires that the investigator deal with the data as given. He may not read into the stories anything which is not there, interpreting them in terms of preconceived notions and symbolic systems that may or may not be in conformity with reality. In Sequential Analysis the analyst deals strictly with the import of the stories as given, with no subjective interpretation of his own. Thus the subject is allowed to speak for himself and in so doing unconsciously reveals his values, principles, and attitudes, in short, the mental set with which he habitually approaches his problems. Since it was hypothesized that the differences between successful and unsuccessful teachers can be found in this area of underlying motivational forces, provided they can be objectively ascertained, this method of analysis was selected as most appropriate.

The Subjects of the Investigation

The Teachers

The sample of teachers was chosen from a Catholic Community of Sisters who have 250 teachers in twenty secondary schools, and approximately 1200 teachers in one hundred thirty-five elementary schools throughout Wisconsin, Illinois, Michigan, Indiana, and Ohio.

To economize on time and expense it was decided to concentrate on large schools and on those areas where a number of smaller schools lie in
close proximity to one another. That the sample might not be biased through selective withdrawal, it was also decided that only those schools would be included in the study whose entire faculty was willing to participate. Of the schools contacted, one did not take part because of unfavorable external circumstances. In all the schools participating, only two teachers refused to cooperate. Since both of these were members of large high school faculties, it was felt that their defection would not bias the sample appreciably, and the schools were included. Eight teachers who were scheduled to take part were dropped because of illness or other unavoidable absence at the time of administration of the tests.

It was felt that the sample thus procured possessed a high degree of randomization. Because of the placement methods used in a Community of Sisters, any teacher may be stationed at any of the schools staffed by the Community. It was just by chance that these particular teachers at this time belonged to the faculty of one of the participating schools.

The sample, as finally constituted, consisted of three hundred Sisters, of whom 150 were teaching in grades four through eight and 150 in grades nine through twelve. Fourteen high schools and twenty-three grade schools were represented, and these were located in four different states.

The teachers ranged in age from twenty to sixty-eight years, and from one to forty-eight years in teaching experience. The median age of the elementary group was 41.5; and that of the secondary group was 47.5. (Table 1) The schools from which these teachers came varied in size, the number of faculty members tested ranging from three to twenty-seven in the high schools, and from one to twelve in the grade schools. In every case but
two, the number represents the entire available number of teachers in the grades covered by the study.

TABLE I

AGE, TEACHING EXPERIENCE, AND TEACHING LEVELS OF TEACHERS PARTICIPATING IN THE STUDY

<table>
<thead>
<tr>
<th>Grades</th>
<th>N</th>
<th>Age Range</th>
<th>Median Age</th>
<th>Years of Experience</th>
</tr>
</thead>
<tbody>
<tr>
<td>4-8</td>
<td>150</td>
<td>20-68</td>
<td>41.5</td>
<td>1-48</td>
</tr>
<tr>
<td>9-12</td>
<td>150</td>
<td>22-68</td>
<td>47.5</td>
<td>1-48</td>
</tr>
</tbody>
</table>

The Pupil Raters

The raters were a total of 10,720 pupils, 6,226 in grades four through eight and 4,494 in grades nine through twelve. They included both boys and girls, and, with the exception of chance absentees, the entire class of each teacher rated. Class size ranged from thirteen to forty-seven in the secondary schools, the average being 30.0, and from twenty-two to fifty-six in the elementary, with an average of 41.5. (Tables 2 and 3)

TABLE II

NUMBER OF PUPILS WHO RATED THEIR TEACHERS

<table>
<thead>
<tr>
<th>Grades</th>
<th>N</th>
<th>Range of Class Size</th>
<th>Mean Class Size</th>
</tr>
</thead>
<tbody>
<tr>
<td>4-8</td>
<td>6226</td>
<td>22-56</td>
<td>41.5</td>
</tr>
<tr>
<td>9-12</td>
<td>4494</td>
<td>13-47</td>
<td>30.0</td>
</tr>
<tr>
<td>All</td>
<td>10720</td>
<td>13-56</td>
<td>35.7</td>
</tr>
</tbody>
</table>
TABLE III

DISTRIBUTION OF TEACHERS AND PUPILS BY GRADE GROUPS

<table>
<thead>
<tr>
<th>Grades</th>
<th>N of Teachers</th>
<th>N of Pupils</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>26</td>
<td>1109</td>
</tr>
<tr>
<td>5-6</td>
<td>49</td>
<td>2034</td>
</tr>
<tr>
<td>7-8</td>
<td>75</td>
<td>3063</td>
</tr>
<tr>
<td>9-10</td>
<td>78</td>
<td>2365</td>
</tr>
<tr>
<td>11-12</td>
<td>72</td>
<td>2129</td>
</tr>
<tr>
<td>All</td>
<td>300</td>
<td>10720</td>
</tr>
</tbody>
</table>

Procedures in Collecting the Data

All the data were collected during the second semester of the school year; thus the pupils had been with the teacher whom they were rating for at least five months.

Administration of the Tests to the Teachers

Each teacher wrote the tests in two sessions. The TAT was given to groups ranging from two to twenty-three according to circumstances. Since the method of TAT analysis used in this study is based on story plot and outcome, not on incidental embellishments of plot or behavior of the subject, group administration is acceptable. The standard directions were given, with emphasis on writing a story that contained a plot and an outcome. There was no time limit; the picture cards were numbered and displayed about
the room, and the stories were written in an unbroken period that varied, according to individual needs, from 55 to 95 minutes. At the second session the Otis Test was administered according to directions, with the one-half hour time limit, followed by the MTAI, also according to standard directions.

The following data were collected from each teacher: age, number of years of teaching experience, and highest degree. The papers were identified only by a number drawn by the teacher and unknown to the investigator.

Administration of the Rating Scale to the Pupils

The rating scale was administered by the investigator to each class in its own classroom. The teacher had previously informed the class that someone was coming to give them a questionnaire to answer, assuring them that neither she nor they would be identified and telling them the number by which she would be designated. She had also previously divided her class into thirds according to achievement, and without revealing the nature of the division had assigned each pupil one of three "code letters" which would indicate to the investigator in which third the respective pupil ranked. The teacher was not in the room during the time the rating scale was administered.

The experimenter informed the class of the purpose and the nature of the study, again assured them that both they and their teacher would be unknown, and promised them that their answers would not be seen by their teacher. The pupils were encouraged to be sincere and to feel completely unrestricted in their answers to the questionnaire. They were asked to mark
on their specially prepared answer sheets their grade (and subject, if in high school), their sex, and the "code letter" which their teacher had given them.

The second part of the scale, the Diagnostic Scale, was answered first. Here the pupils marked each one of the forty-nine items that they considered true of their teacher. In the fourth and fifth grades the investigator avoided reading problems by reading the scale aloud as the pupils followed silently. In all grades any words that gave difficulty were explained upon request.

After the forty-nine items had been checked, the pupils were asked to arrange the seven areas into which these statements are divided in the order in which they thought their teacher "was good in them." They were told to mark with "1" her best area, with "2" her second best, and so on. They were also permitted to list any other "good" or "bad" qualities of their teacher that were not included in the check list.

By this time it was felt that the pupils had a fairly clear understanding of what each area implied and were now ready to do Part I of the Scale, the Area Scale. It was this part of the instrument that was being considered primarily for use in the investigation, while the preceding steps were meant to prepare the raters so as to increase the validity of the scale.

The Area Scale consists of seven questions corresponding to the seven divisions of items on the Diagnostic Scale. In the administration of the Area Scale the investigator did not follow the given directions, which call for a rating on a five-point scale. In place of the printed instructions, the pupils were directed as follows:
For each question compare this teacher with all the other teachers you have ever had. The first question says, "How well do you like your teacher?" If you like Sister N. better than any other teacher you have ever had, put a ring around number 5; if you like Sister N. the least of all the teachers you have ever had, put a ring around number 1; if you like Sister N. very much, but you have had one teacher that you liked better, put a ring around 4; if you do not like Sister N. very much and have had only one teacher that you liked less, put a ring around 2; if you think Sister N. is just about in the middle of all your teachers as far as your liking is concerned, put a ring around 3. Now do the same for all the other questions; for each one, compare Sister N. with all the other teachers you have ever had.

In the fourth grade rooms the above instruction was repeated for all the questions. In the fifth and sixth grades it was repeated for the second and third questions; and in the seventh, eighth, and ninth grades, for the second only. With the older pupils it was not repeated. Also, the directions with regard to encircling numbers 2 and 4 varied according to the age of the pupils. For grades four, five, and six, they were given as above. In grades seven, eight, and nine, the pupils were told, "If you had one or two teachers that you liked better, put a ring around 4; and if you had one or two that you liked less, encircle 2." In grades ten, eleven, and twelve this was changed to "one, two, or three." This was, of course, to take care of the greater number of teachers previously had by the older pupils and to prevent too much clustering at the middle of the scale with too little discrimination.

These directions strengthened the scale considerably by giving both the pupils and the investigator a more tangible basis for comparison. The numbers thus indicated rank positions and could be considered arbitrary weights or scores for purposes of calculation; that is, each teacher who was ranked at the top of a particular pupil's teachers received a score of
five for the question on which she was so ranked. Understanding these numbers as scores, the experimenter could more intelligibly manipulate them statistically.

The pupils' answer sheets were not handled by the investigator in the classroom. When the rating was finished, the pupils were asked to write their teacher's identification number on their papers. The latter were then collected by one of the pupils who put them into an envelope, sealed it, and carried it to a designated place, usually the school office, where it lost its identity among the envelopes of other classes.
CHAPTER III

ANALYSIS OF THE DATA

The present chapter will describe the analytical procedures used in handling the data and the results obtained in this analysis. The treatment of the pupils' ratings and the setting up of criterion groups on the basis of those ratings will be discussed first. In the second part of the chapter the results of the intelligence test and the attitude inventory will be described with reference to the entire sample of three hundred teachers. This will be followed by an account of the investigation with the paired samples, using one hundred cases and involving the MTAI and the TAT.

Analysis of the Rating Scale

In the setting up of the criterion groups only the Area Scale was used. A pilot study had indicated that there was no significant difference between the ratings of the boys and the girls and between the ratings of the different achievement groups. No further distinction was made, therefore, between the ratings of these groups.

Each teacher's mean score on each of the seven questions was computed, using the designated figure in the five-point scale as a score. Inspection revealed that these means became progressively lower from the fourth grade to the twelfth, indicating that the pupils became more critical and discriminating as they became older. To have a common basis for comparison,
therefore, centile norms were set up in each of the seven areas for each of the following five grade groups: 4, 5-6, 7-8, 9-10, 11-12. The smoothed ogive was used in setting up these centile norms and each teacher's rank was read from the ogive.

The tendency of the older children to rate their teachers lower is evident in Figure 1, which shows the centile equivalents of a score of 4.0 for each of these groups on each question. Thus, 87% of the 11-12 grade teachers rated below a mean score of 4.0 on Question 1, while only 28% of the fourth grade teachers rated below that score on the same question.

![Graph showing centile equivalents for various grade levels.](image-url)

**Fig. 1. CENTILE EQUIVALENTS OF AN AVERAGE RATING OF 4.0 FOR TEACHERS ON VARIOUS GRADE LEVELS**
Intercorrelations among the seven ratings, as indicated by phi coefficients, were relatively high both for the grade school and for the high school, ranging from .14 to .73. The phi coefficients were computed on the proportion of teachers above and below the 50th centile on each rating. In the grade school all intercorrelations, and in the high school all but five, were significant beyond the 1% level. Questions IV and VI, "How fair is your teacher in grading" and "Does your teacher give the children the right amount of work to do," showed the least amount of relationship to the other questions of the scale. (Tables IV and V)

TABLE IV
INTERCORRELATIONS (PHI COEFFICIENTS)
OF THE SEVEN RATINGS
(High School)
N = 150

<table>
<thead>
<tr>
<th></th>
<th>II</th>
<th>III</th>
<th>IV</th>
<th>V</th>
<th>VI</th>
<th>VII</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>.46</td>
<td>.69</td>
<td>.31</td>
<td>.29</td>
<td>.31</td>
<td>.63</td>
</tr>
<tr>
<td>II</td>
<td>.27</td>
<td>.20</td>
<td>.49</td>
<td>.31</td>
<td>.49</td>
<td></td>
</tr>
<tr>
<td>III</td>
<td>.27</td>
<td>.16</td>
<td>.18</td>
<td>.40</td>
<td></td>
<td></td>
</tr>
<tr>
<td>IV</td>
<td>.17</td>
<td>.14</td>
<td>.28</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>V</td>
<td>.22</td>
<td>.30</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>VI</td>
<td>.44</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Significance: Above .21 beyond 1% level; .20, 2%; .16 to .18, 5%; .14 NS

The high correlations suggest a good deal of halo effect and would seem to indicate that a teacher was probably evaluated as a whole on the Area Scale. Since the purpose of the rating in the study was to arrive at
criterion groups that were clearly distinguished as such in the minds of their pupils and not to discriminate among the various characteristics of the teachers, this intercorrelation was not considered a disadvantage. However, the relative amounts of correlation might help to reveal which characteristics seemed to have influenced the pupils most in their rating.

TABLE V

INTERCORRELATIONS (PHI COEFFICIENTS)
OF THE SEVEN RATINGS
(Grade School)
N = 150

<table>
<thead>
<tr>
<th></th>
<th>II</th>
<th>III</th>
<th>IV</th>
<th>V</th>
<th>VI</th>
<th>VII</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>.57</td>
<td>.68</td>
<td>.49</td>
<td>.52</td>
<td>.44</td>
<td>.65</td>
</tr>
<tr>
<td>II</td>
<td>.44</td>
<td>.36</td>
<td>.49</td>
<td>.36</td>
<td>.62</td>
<td></td>
</tr>
<tr>
<td>III</td>
<td>.38</td>
<td>.48</td>
<td>.44</td>
<td>.73</td>
<td></td>
<td></td>
</tr>
<tr>
<td>IV</td>
<td>.52</td>
<td>.44</td>
<td>.41</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>V</td>
<td>.44</td>
<td>.52</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>VI</td>
<td></td>
<td>.57</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

All coefficients significant beyond the 1% level

In both grade and high school there was a much closer relationship between Question I, "How well do you like this teacher," and Questions II, III, and VII than there was between Question I and the other three questions. This would seem to indicate that the teacher's kindness and friendliness (III), the way she taught her lessons (VII), and her ability to explain (II), had much more bearing on how well she was liked, than had her fairness in marking (IV), her discipline (V), and the amount of work she gave (VI).
In other words, a teacher may be very well liked even though her marking, discipline, and assignments may meet with considerable disapproval, and vice versa. This was particularly true in the high school, though the pattern was identical on all levels. The grade school pupils' liking for the teacher was somewhat more related to the way the teacher disciplines than was the high school pupils'. In the grade school the highest correlation was between the teacher's kindness and understanding and the pupils' liking for their lessons, which would seem to imply that the elementary school pupil tends to consider a lesson "good" when his difficulties are seen and provided for with sympathy.

**Analysis of the Data from the Whole Sample**

The entire sample of three hundred teachers was used in the first part of the experiment to discover whether there were any significant differences between groups in intelligence and expressed teacher attitudes.

**The Intelligence Test**

It was found that the median Gamma IQ of the total sample of three hundred teachers was 109.5. The median of the grade school group was 108.2; while that of the high school group was 110.6 (Table VI) To determine whether there was any relationship between the intelligence of the teachers as measured and the way they were rated in each of the seven areas,

---

1Gamma IQ's above 100 are generally lower than IQ's derived from other measures. This is particularly true on the higher levels. The highest possible IQ on the Otis Test is 138.
the hypothesis of no difference was tested by chi square. A 2 x 2 contingency table was set up for the high school teachers and the grade school teachers separately for each of the seven questions. The categories used were number of teachers above and below the midpoint of the rating distribution and the number having IQ's above and below the median of their group.

**TABLE VI**

**MEASURES OF CENTRAL TENDENCY AND VARIABILITY OF INTELLIGENCE TEST SCORES OF THE ENTIRE SAMPLE**

<table>
<thead>
<tr>
<th>Statistic</th>
<th>Elementary teachers</th>
<th>Secondary teachers</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>150</td>
<td>150</td>
</tr>
<tr>
<td>M</td>
<td>107.6 ± 1.03</td>
<td>110.3 ± 0.94</td>
</tr>
<tr>
<td>σ</td>
<td>12.6 ± 0.73</td>
<td>11.5 ± 0.66</td>
</tr>
<tr>
<td>Mdn</td>
<td>108.2</td>
<td>110.6</td>
</tr>
</tbody>
</table>

Chi square tests revealed that there was a significant relationship between intelligence test scores of the high school teachers and Question II: "How clearly does your teacher explain things," with the more intelligent teachers rated higher. Chi square was significant beyond the 1% level on this question, as well as on the sixth: "Does your teacher give the children the right amount of work to do?" These relationships were further defined in terms of the biserial r, which was found to be .44 ± .09 between intelligence scores and Question II, and .33 ± .09 between intelligence and Question VI.

The chi square test also showed a relationship between intelligence
and Question III: "How kind, friendly, and understanding is your teacher," that was significant at the 5% level. The biserial $r (0.15 \pm 0.10)$, however, was not significant. The secondary teachers with higher intelligence scores were also rated higher on the other four questions, though not significantly so. The detailed picture of the relationship between the intelligence test scores of the high school teachers and their rating on each of the seven questions is shown in Table VII.

**TABLE VII**

**RELATIONSHIP BETWEEN INTELLIGENCE SCORES OF HIGH SCHOOL TEACHERS AND THEIR RATING ON EACH OF THE SEVEN QUESTIONS**

<table>
<thead>
<tr>
<th>Questions of rating scale</th>
<th>R</th>
<th>Frequencies</th>
<th>Gamma IQ</th>
<th>Chi square</th>
<th>Level of sig.</th>
<th>$r_{bis}$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>111+</td>
<td>110-</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1. How well do you like your teacher?</td>
<td>50+</td>
<td>43</td>
<td>35</td>
<td>1.709</td>
<td>NS</td>
<td></td>
</tr>
<tr>
<td></td>
<td>49-</td>
<td>32</td>
<td>40</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. How clearly can your teacher explain things?</td>
<td>50+</td>
<td>47</td>
<td>28</td>
<td>8.64</td>
<td>S,1%</td>
<td>0.44+0.09</td>
</tr>
<tr>
<td></td>
<td>49-</td>
<td>28</td>
<td>47</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. How kind, friendly, and understanding is your teacher?</td>
<td>50+</td>
<td>47</td>
<td>34</td>
<td>4.536</td>
<td>NS</td>
<td>0.15+0.10</td>
</tr>
<tr>
<td></td>
<td>49-</td>
<td>28</td>
<td>41</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. How fair is your teacher in grading?</td>
<td>50+</td>
<td>41</td>
<td>34</td>
<td>1.307</td>
<td>NS</td>
<td></td>
</tr>
<tr>
<td></td>
<td>49-</td>
<td>34</td>
<td>41</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5. How well does your teacher keep order with the children?</td>
<td>50+</td>
<td>44</td>
<td>36</td>
<td>1.714</td>
<td>NS</td>
<td></td>
</tr>
<tr>
<td></td>
<td>49-</td>
<td>31</td>
<td>39</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6. Does your teacher give the right amount of work to do?</td>
<td>50+</td>
<td>47</td>
<td>30</td>
<td>7.712</td>
<td>S,1%</td>
<td>0.33+0.09</td>
</tr>
<tr>
<td></td>
<td>49-</td>
<td>28</td>
<td>45</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7. How well do you like the lessons taught by this teacher?</td>
<td>50+</td>
<td>40</td>
<td>36</td>
<td>.427</td>
<td>NS</td>
<td></td>
</tr>
<tr>
<td></td>
<td>49-</td>
<td>35</td>
<td>39</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

$N = 150$  \quad R = $rating in centiles$
These results may be interpreted to mean that the more intelligent high school teacher explains more clearly, and that her superiority in this respect is recognized by her pupils. She probably has a better command of her subject and is thus able both to explain well and to make more judicious assignments. A teacher with lower intelligence is, however, quite as capable apparently of winning the affection of her pupils, of keeping order, and of grading fairly. Liking for the subject matter may have strongly influenced the replies to Question VII, so that pupils may have indicated a liking for the lessons on that basis rather than on the ability of the teacher to make the lessons interesting.

In the elementary group there were no significant relationships between intelligence test scores and ratings on any of the questions, though the scores of the high-rated groups were consistently somewhat higher. Neither were there any significant differences in intelligence between the high-rated elementary and the high-rated secondary school teachers.

Analysis of the Attitude Inventory

The MTAI was scored according to the standard scoring stencil prepared by the authors. Since inspection revealed some differences in the way teachers of different grade levels tended to score on the Inventory, each of the grade groups, as originally set up in terms of their rating, was studied separately.

The criterion groups were determined on the basis of the teachers' rating on Question I: "How well do you like this teacher?" This was done because that question correlated most highly with all the other questions
and seemed to correspond most closely with the criterion on which the MTAI
was originally validated.

Mean MTAI scores were computed for the 20 teachers rated highest and
the 20 rated lowest in each of the five groups, and a test of significance
was applied to the differences between means. Sizable differences were
found only in the 7-8 and 9-10 grade groups, where the teachers who were
rated high also made higher scores on the MTAI. Only in the first of these,
however, was the difference significant at the 5\% level. Table VIII shows
that on two grade levels, namely, 7 and 11-12, the means of the high groups
were actually lower than those of the low groups. (The range of scores on
the MTAI is from +150 to -150, the higher scores indicating the more favor-
able attitudes.)

**TABLE VIII**

<table>
<thead>
<tr>
<th>Statistic</th>
<th>Grade 4</th>
<th>Grades 5-6</th>
<th>Grades 7-8</th>
<th>Grades 9-10</th>
<th>Grades 11-12</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>H</td>
<td>L</td>
<td>H</td>
<td>L</td>
<td>H</td>
</tr>
<tr>
<td>N</td>
<td>10*</td>
<td>10*</td>
<td>20</td>
<td>20</td>
<td>20</td>
</tr>
<tr>
<td>M</td>
<td>-19</td>
<td>-3</td>
<td>1</td>
<td>3</td>
<td>18</td>
</tr>
<tr>
<td>σ</td>
<td>31.5</td>
<td>41.4</td>
<td>32.9</td>
<td>35.3</td>
<td>30.5</td>
</tr>
<tr>
<td>σ&lt;sub&gt;M&lt;/sub&gt;</td>
<td>10.5</td>
<td>13.7</td>
<td>7.6</td>
<td>8.1</td>
<td>7.0</td>
</tr>
<tr>
<td>D&lt;sub&gt;M&lt;/sub&gt;</td>
<td>-16</td>
<td>4</td>
<td>24</td>
<td>14</td>
<td>14</td>
</tr>
<tr>
<td>σ&lt;sub&gt;dM&lt;/sub&gt;</td>
<td>17.2</td>
<td>11.1</td>
<td>10.3</td>
<td>10.9</td>
<td>9.9</td>
</tr>
<tr>
<td>t</td>
<td>.93</td>
<td>.36</td>
<td>2.33</td>
<td>1.28</td>
<td>.10</td>
</tr>
<tr>
<td>Sig.</td>
<td>NS</td>
<td>NS</td>
<td>8.5%</td>
<td>NS</td>
<td>NS</td>
</tr>
</tbody>
</table>

*<sup>*N = 10, because of small number of fourth grade teachers.**
In grades 7-8 a mean score of 18 for the high group and -6 for the low group resulted in a $t$ of 2.33, where 2.025 is required for significance at the 5% level, df being 38. This would seem to indicate that the pupils of grades 7-8 are more sensitive to, or more influenced in their liking by, the attitudes of their teachers than are the pupils of other grades.

The second step in the experiment with the MTAI was to determine whether there were any differences in the way the high and low groups at the various grade levels answered certain types of questions. The 150 items of the inventory were classified according to their content into the following categories:

(1) Attitudes toward children and teaching; this was subdivided as follows:

a. Teacher's opinion of pupil conduct (14 items)
Examples: 1. Most children are obedient.
113. Pupils like to annoy the teacher.

b. Teacher's liking for pupils and teaching (16 items)
Examples: 5. Teaching never gets monotonous.
98. Pupils can be very boring at times.

c. Teacher's opinion of pupils' ability and effort (20 items)
Examples: 19. Pupils have it too easy in the modern school.
23. Most pupils do not make an adequate effort to prepare their lessons.

(2) Attitudes toward discipline; this was subdivided:

a. Teacher's opinion of present-day conditions (13 items)
Examples: 24. Too many children nowadays are allowed to have their own way.
51. Discipline problems are the teacher's greatest problem.

b. Teacher's opinion of what should be done about discipline (26 items)

---

2 The formula used was that given by Fisher for small samples of equal size.
Examples: 3. Minor disciplinary situations should sometimes be turned into jokes.

39. To maintain good discipline in the classroom a teacher needs to be "hard-boiled."

(3) Teacher's knowledge of psychological and sociological backgrounds (25 items)

Examples: 28. The boastful child is usually over-confident of his ability.

49. A teacher should not be expected to be sympathetic toward truants.

(4) Teacher's understanding of principles of teaching (25 items)

Examples: 16. A pupil's failure is seldom the fault of the teacher.

37. Standards of work should vary with the pupil.

(5) Teacher's attitude toward authority (11 items)

Examples: 11. Unquestioning obedience in a child is not desirable.

144. Teachers can be in the wrong as well as pupils.

Still using Question I of the rating scale as criterion, the high and low teachers in each grade group were scored in each of the above categories. The mean was computed for each group and the t test of significance was applied to the difference between means. The results are presented in Tables IX to XIII. In grades 4, 5-6, and 11-12 there were no significant differences in any of the categories. The 9-10 grade teachers differed in two categories only, l-c and 2-b, in both of which the high-rated teachers made higher scores. The differences were significant at the 5% level, t being 2.03 for Category l-c and 2.14 for Category 2-b.

In the 7-8 grade group Categories l-a, l-b, and 2-b showed differences, also in favor of the high-rated group. In the first, t was 2.69, which is just short of being significant at the 1% level, 2.91 being required for this level of significance, df, 38. In Category l-b, t was 3.33, which is significant beyond the 1% level. Category 2-b was significant at the 5% level, t being 2.54.
### TABLE IX
DIFFERENCES BETWEEN HIGH AND LOW GROUPS
ON CATEGORIES OF THE MTAI

**Grade 4**

<table>
<thead>
<tr>
<th>Statistic</th>
<th>1-a</th>
<th>1-b</th>
<th>1-c</th>
<th>2-a</th>
<th>2-b</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>M (High)</td>
<td>-1.3</td>
<td>-0.5</td>
<td>+1.4</td>
<td>-0.6</td>
<td>-10.9</td>
<td>0</td>
<td>-0.6</td>
<td>-2.4</td>
</tr>
<tr>
<td>M (Low)</td>
<td>-0.2</td>
<td>+0.4</td>
<td>+1.6</td>
<td>+0.1</td>
<td>-5.6</td>
<td>-1</td>
<td>-0.3</td>
<td>-4.7</td>
</tr>
<tr>
<td>$D_M$</td>
<td>-1.1</td>
<td>-0.9</td>
<td>-0.2</td>
<td>-0.7</td>
<td>-5.3</td>
<td>+1</td>
<td>-0.3</td>
<td>+2.3</td>
</tr>
<tr>
<td>$\sigma_{dM}$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$t$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sig.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>NS</td>
<td></td>
</tr>
</tbody>
</table>

$N = 10$ high, $10$ low

### TABLE X
DIFFERENCES BETWEEN HIGH AND LOW GROUPS
ON CATEGORIES OF THE MTAI

**Grades 5-6**

<table>
<thead>
<tr>
<th>Statistic</th>
<th>1-a</th>
<th>1-b</th>
<th>1-c</th>
<th>2-a</th>
<th>2-b</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>M (High)</td>
<td>-2.1</td>
<td>-0.8</td>
<td>+2.8</td>
<td>-0.4</td>
<td>-2.2</td>
<td>+0.6</td>
<td>+2.0</td>
<td>-1.4</td>
</tr>
<tr>
<td>M (Low)</td>
<td>+0.3</td>
<td>-2.1</td>
<td>+4.7</td>
<td>+0.1</td>
<td>-5.7</td>
<td>+0.8</td>
<td>-1.0</td>
<td>-1.2</td>
</tr>
<tr>
<td>$D_M$</td>
<td>-2.4</td>
<td>+1.3</td>
<td>-1.9</td>
<td>-0.5</td>
<td>+3.5</td>
<td>-0.2</td>
<td>+3.0</td>
<td>-0.2</td>
</tr>
<tr>
<td>$\sigma_{dM}$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$t$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1.41</td>
<td></td>
</tr>
<tr>
<td>Sig.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>NS</td>
<td></td>
</tr>
</tbody>
</table>

$N = 20$ high, $20$ low
### TABLE XI

**DIFFERENCES BETWEEN HIGH AND LOW GROUPS ON CATEGORIES OF THE MTAI**

**Grades 7-8**

<table>
<thead>
<tr>
<th>Statistic</th>
<th>Categories</th>
<th>1-a</th>
<th>1-b</th>
<th>1-c</th>
<th>2-a</th>
<th>2-b</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>M (High)</td>
<td>+1.4</td>
<td>+2.0</td>
<td>+4.4</td>
<td>+0.7</td>
<td>-0.6</td>
<td>+5.0</td>
<td>+3.4</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>M (Low)</td>
<td>-3.1</td>
<td>-2.0</td>
<td>+4.6</td>
<td>-0.2</td>
<td>-6.0</td>
<td>+0.5</td>
<td>0</td>
<td>-1.4</td>
<td></td>
</tr>
<tr>
<td>𝑑_𝑀</td>
<td>+4.5</td>
<td>+4.0</td>
<td>+0.9</td>
<td>+5.4</td>
<td>+4.5</td>
<td>+3.4</td>
<td>+1.4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>𝜎_𝑑_𝑀</td>
<td>1.67</td>
<td>1.2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>𝑡</td>
<td>2.69</td>
<td>3.33</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sig.</td>
<td>2%</td>
<td>1%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

N = 20 high, 20 low

### TABLE XII

**DIFFERENCES BETWEEN HIGH AND LOW GROUPS ON CATEGORIES OF THE MTAI**

**Grades 9-10**

<table>
<thead>
<tr>
<th>Statistic</th>
<th>Categories</th>
<th>1-a</th>
<th>1-b</th>
<th>1-c</th>
<th>2-a</th>
<th>2-b</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>M (High)</td>
<td>+1.5</td>
<td>+0.1</td>
<td>+7.0</td>
<td>+2.6</td>
<td>-1.0</td>
<td>+1.3</td>
<td>+3.9</td>
<td>-0.2</td>
<td></td>
</tr>
<tr>
<td>M (Low)</td>
<td>-0.8</td>
<td>-0.7</td>
<td>+2.0</td>
<td>-1.1</td>
<td>-1.0</td>
<td>+2.3</td>
<td>+0.2</td>
<td>-0.6</td>
<td></td>
</tr>
<tr>
<td>𝑑_𝑀</td>
<td>+2.3</td>
<td>+0.8</td>
<td>+5.0</td>
<td>+3.7</td>
<td>0</td>
<td>-1.0</td>
<td>+2.7</td>
<td>+0.4</td>
<td></td>
</tr>
<tr>
<td>𝜎_𝑑_𝑀</td>
<td>1.69</td>
<td>2.46</td>
<td>1.73</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>𝑡</td>
<td>1.33</td>
<td>2.03</td>
<td>2.14</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sig.</td>
<td>NS</td>
<td>5%</td>
<td>5%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

N = 20 high, 20 low
TABLE XIII
Differences between High and Low Groups
On Categories of the MTAI
Grades 11-12

<table>
<thead>
<tr>
<th>Statistic</th>
<th>Categories</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1-a</td>
</tr>
<tr>
<td><strong>M (High)</strong></td>
<td>+0.3</td>
</tr>
<tr>
<td><strong>M (Low)</strong></td>
<td>-0.8</td>
</tr>
<tr>
<td><strong>D_M</strong></td>
<td>+1.1</td>
</tr>
<tr>
<td><strong>σ_d_M</strong></td>
<td>2.25</td>
</tr>
<tr>
<td><strong>t</strong></td>
<td>1.11</td>
</tr>
<tr>
<td><strong>Sig.</strong></td>
<td>NS</td>
</tr>
</tbody>
</table>

N = 20 high, 20 low

The differences noted earlier in the total means of grades 7-8 and 9-10 are apparently accounted for by Categories 1 and 2. Since these categories have to do with the teacher's attitudes toward children and teaching and toward discipline, it may be concluded that the pupils at this age level are particularly keen in sensing their teacher's evaluation of them and her work, and that they demand the more permissive attitudes toward discipline represented by the higher MTAI scores.

Investigation of the Paired Samples

The work done up to this point suggested strongly that there were differences between the older and the younger teachers in their scoring on both the intelligence test and the attitude inventory, and that age might also have been a factor in the way they were rated by their pupils. Chi
square tests showed that this supposition was definitely true with regard to the intelligence test scores. The median age of the grade school teachers was 41.5; that of the high school teachers was 47.5. When the teachers were categorized on the basis of median age and median IQ, it was found that the younger teachers made significantly higher intelligence scores than the older teachers.\(^3\) This was true in both grade and high schools, chi square being 56.467 and 16.669 respectively, where only 6.635 is required for significance at the 1\(^\circ\) level. (Table XIV)

On the MTAI the median scores of the elementary and the secondary teachers were +7.5 and +4.5 respectively. The hypothesis of a difference was supported only with regard to the grade school, where the younger teachers scored significantly higher than the older ones, chi square being 9.626. There was no significant difference between the MTAI scores of the younger and the older teachers in the secondary schools, though there was a strong tendency for the younger ones to score higher.

Chi square tests showed that age of the teacher was no factor in the rating of the grade school pupils, whereas it was an important factor in the high school. Here chi square was 11.257, significant beyond the 1\(^\circ\) level, with more of the younger teachers rated above the 50th centile on 5 out of the 7 questions. The interrelationships of these factors in the two groups are shown in Tables XIV and XV.

\(^3\)This difference must be attributed, to an unknown degree, of course, to the fact that the younger teachers were much more "test wise" than the older ones, for many of whom this was a first experience.
### TABLE XIV

**RELATIONSHIP OF AGE TO INTELLIGENCE SCORES, MTAI SCORES, AND RATING**

**Grades 4-8**

<table>
<thead>
<tr>
<th>Measure</th>
<th>Category</th>
<th>Age</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Above 41.5</td>
<td>Below 41.5</td>
<td>Chi square</td>
<td>Sig.</td>
<td></td>
</tr>
<tr>
<td>Gamma IQ</td>
<td>Above 108.2</td>
<td>13</td>
<td>60</td>
<td>56.467</td>
<td>.1%</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Below 108.2</td>
<td>62</td>
<td>15</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MTAI</td>
<td>Above +7.5</td>
<td>28</td>
<td>47</td>
<td>9.626</td>
<td>.1%</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Below +7.5</td>
<td>47</td>
<td>28</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rating*</td>
<td>Above median</td>
<td>22</td>
<td>18</td>
<td>.046</td>
<td>NS</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Below median</td>
<td>27</td>
<td>22</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Rating is based on 5 out of the 7 questions

### TABLE XV

**RELATIONSHIP OF AGE TO INTELLIGENCE SCORES, MTAI SCORES, AND RATING**

**Grades 9-12**

<table>
<thead>
<tr>
<th>Measure</th>
<th>Category</th>
<th>Age</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Above 47.5</td>
<td>Below 47.5</td>
<td>Chi square</td>
<td>Sig.</td>
<td></td>
</tr>
<tr>
<td>Gamma IQ</td>
<td>Above 110.6</td>
<td>25</td>
<td>51</td>
<td>16.667</td>
<td>.1%</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Below 110.6</td>
<td>50</td>
<td>24</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MTAI</td>
<td>Above +4.5</td>
<td>35</td>
<td>40</td>
<td>.667</td>
<td>NS</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Below +4.5</td>
<td>40</td>
<td>35</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rating*</td>
<td>Above median</td>
<td>13</td>
<td>28</td>
<td>11.253</td>
<td>.1%</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Below median</td>
<td>27</td>
<td>12</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Rating is based on 5 out of the 7 questions
The Paired Samples

Because of these findings it was decided to do all further experimentation with two samples of 50 teachers each, paired on the basis of intelligence scores and age, and rating at the extremes in five out of the seven questions of the rating scale. For this purpose Questions IV and VI, which correlated least with the rest of the scale, and which were most difficult to interpret in view of varying regulations with regard to grading and assigning of work, were dropped. The final sample was made up of 25 pairs of elementary teachers and 25 pairs of secondary teachers, the high groups, with rare exceptions, rating above the 50th centile on the five remaining questions and the low groups rating below.

In addition to age and intelligence several other factors were taken into consideration in the pairing of the samples. To offset any differences in rating that might exist between different kinds of schools, the pairs were, wherever possible, taken from the same school, or from the same type of school, as regards size, location, and population. Thus, a teacher in a coeducational school was not paired with one in an all-girl school, and a teacher in a large city school was not paired with a teacher in a small town school. All pairs came from the same grade group, and in the high schools, from the same subject or kind of subject, if possible. Care was taken not to pair a teacher of a general education subject, or one required of all pupils, with a teacher of a special subject or an elective.

The high and low samples, thus constituted, appeared to be as alike as possible in all the controllable variables and as different as possible in the criterion. Tables XVI, XVII, and XVIII show the comparative statistics
of elementary and secondary high and low groups on age, intelligence scores, and rating by the pupils. The high and low groups had approximately equal means, standard deviations, and skewness. The greatest divergence was in intelligence scores of the secondary teachers, where the difference in means was, however, in favor of the low group and was not statistically significant. 4

**TABLE XVI**

COMPARATIVE STATISTICS OF THE PAIRED GROUPS ON AGE

<table>
<thead>
<tr>
<th>Statistic</th>
<th>Elementary teachers</th>
<th>Secondary teachers</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>High-rated</td>
<td>Low-rated</td>
</tr>
<tr>
<td>N</td>
<td>25</td>
<td>25</td>
</tr>
<tr>
<td>MdN</td>
<td>36</td>
<td>37</td>
</tr>
<tr>
<td>M</td>
<td>37.7 ± 2.0</td>
<td>38.7 ± 1.9</td>
</tr>
<tr>
<td>σ</td>
<td>9.9 ± 1.4</td>
<td>9.3 ± 1.3</td>
</tr>
<tr>
<td>Dₘ</td>
<td>1.0</td>
<td></td>
</tr>
<tr>
<td>σₘ</td>
<td>0.989</td>
<td></td>
</tr>
<tr>
<td>t</td>
<td>1.01</td>
<td></td>
</tr>
</tbody>
</table>

In the computation of the median ratings, all the ratings within a group were pooled, each group having a total of 125 ratings; that is, five ratings for each of 25 teachers. In both low groups the median rating was below the 25th centile of the total distribution, and that of the high

4 According to Fisher's formula for differences between correlated pairs of means.
groups was well above the 75th centile. This indicated that the groups were clearly at the extremes.

**TABLE XVII**

**COMPARATIVE STATISTICS OF THE PAIRED GROUPS ON INTELLIGENCE SCORES**

<table>
<thead>
<tr>
<th>Statistic</th>
<th>Elementary teachers</th>
<th>Secondary teachers</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>High-rated</td>
<td>Low-rated</td>
</tr>
<tr>
<td>N</td>
<td>25</td>
<td>25</td>
</tr>
<tr>
<td>Mdn</td>
<td>115</td>
<td>114</td>
</tr>
<tr>
<td>M</td>
<td>113.9 ± 2.3</td>
<td>113.5 ± 2.0</td>
</tr>
<tr>
<td>σ</td>
<td>11.1 ± 1.6</td>
<td>9.7 ± 1.4</td>
</tr>
<tr>
<td>D_M</td>
<td>.4</td>
<td></td>
</tr>
<tr>
<td>σ_dM</td>
<td></td>
<td>1.7</td>
</tr>
<tr>
<td>t</td>
<td></td>
<td>.24</td>
</tr>
</tbody>
</table>

**TABLE XVIII**

**COMPARATIVE STATISTICS OF THE PAIRED GROUPS ON RATINGS**

<table>
<thead>
<tr>
<th>Statistic</th>
<th>Elementary teachers</th>
<th>Secondary teachers</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>High-rated</td>
<td>Low-rated</td>
</tr>
<tr>
<td>N*</td>
<td>125</td>
<td>125</td>
</tr>
<tr>
<td>Mdn</td>
<td>86.7</td>
<td>16.3</td>
</tr>
<tr>
<td>Q</td>
<td>7.0</td>
<td>9.1</td>
</tr>
</tbody>
</table>

*Five ratings for each of 25 teachers*
The MTAI and the Paired Samples

To determine whether there were significant differences on the MTAI between the teachers in the paired samples, the $t$ test was applied to the difference between means. Again, there were no significant differences in the high school groups. The means were $10 \pm 7.1$ and $8.7 \pm 6.1$ respectively for high and low groups, $t$ of .15 being insignificant. In the elementary school, however, means of $21 \pm 5.4$ for the high group and $3 \pm 6.8$ for the low, produced a $t$ of $2.40$, which is significant beyond the $5\%$ level, $df$ being $24.5$.

Another attempt was made to locate these differences with respect to the categories set up earlier. Computations were made only for those categories in which there were appreciable differences. Again there were no significant differences in the high school groups, though Category 2-a, which has to do with the teacher's opinion of present-day conditions with regard to discipline, was just short of being significant at the $5\%$ level, with a $t$ of $2.06$. The results of the MTAI study in the high school groups are shown in Table XIX.

In the grade school five categories were tested, of which three proved significantly different, the high-rated teachers making the higher scores. Category 1-b, which has to do with the teacher's liking of pupils and teaching, revealed the greatest differences between high and low groups, $t$ being $4.55$, which is significant beyond the $0.1\%$ level. Category 2-b, dealing with what should be done about discipline, was the second most significant.

---

$^5$According to Fisher's formula for paired samples.
area with a $t$ of 2.76, which is between the 1% and 2% levels of confidence; while differences in Category 1-a, which expresses the teacher's opinion of pupils' conduct, were significant beyond the 2% level. (Table XX)

**TABLE XIX**

**DIFFERENCES ON THE MTAI BETWEEN PAIRED SECONDARY GROUPS**

<table>
<thead>
<tr>
<th>Statistic</th>
<th>Categories of MTAI</th>
<th>Total scores</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1-c</td>
<td>2-a</td>
</tr>
<tr>
<td>$N$ pairs</td>
<td>25</td>
<td>25</td>
</tr>
<tr>
<td>$M_d$</td>
<td>2.8</td>
<td>2.6</td>
</tr>
<tr>
<td>$\sigma_{M_d}$</td>
<td>1.80</td>
<td>1.26</td>
</tr>
<tr>
<td>$t$</td>
<td>1.56</td>
<td>2.06</td>
</tr>
<tr>
<td>Sig.</td>
<td>NS</td>
<td>NS</td>
</tr>
</tbody>
</table>

**TABLE XX**

**DIFFERENCES ON THE MTAI BETWEEN PAIRED ELEMENTARY GROUPS**

<table>
<thead>
<tr>
<th>Statistic</th>
<th>Categories of MTAI</th>
<th>Total scores</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1-a</td>
<td>1-b</td>
</tr>
<tr>
<td>$N$ pairs</td>
<td>25</td>
<td>25</td>
</tr>
<tr>
<td>$M_d$</td>
<td>3.2</td>
<td>3.6</td>
</tr>
<tr>
<td>$\sigma_{M_d}$</td>
<td>1.26</td>
<td>.791</td>
</tr>
<tr>
<td>$t$</td>
<td>2.52</td>
<td>4.55</td>
</tr>
<tr>
<td>Sig.</td>
<td>8, 2%</td>
<td>8, 1%</td>
</tr>
</tbody>
</table>
In a further attempt to distinguish between groups, the MTAI was item-analyzed to determine whether there were any questions on which the high and low teachers were clearly different. Again there were no differences in the high school group. Among the elementary teachers eight items were significantly different at or beyond the 5% level. These items were numbers 5, 101, 39, 38, 54, 89, 145, and 90. Table XXI enumerates these items and indicates the direction in which the differences lay. The answers, SA, A, U, D, SD, as given for the high and low groups, were in only two cases the combination of right and wrong answers as given by the authors of the inventory.

All but one of the above items appear in Category 1, indicating again that the difference between the high and low groups, as distinguishable by the MTAI, lay chiefly in the teacher's attitudes toward children and teaching. The only item not in this category is item number 39, which has to do in a general way with discipline and is closely allied to the others.

The TAT and the Paired Samples

The TAT's of the one hundred cases in the paired samples were analyzed. The first step was TAT Sequential Analysis. Since according to this method the investigator simply restates the import of the story, the results of this analysis were in the form of twelve sequential statements (one for each of twelve stories) for each case. In other words, the investigator now had a sample of twelve statements covering each subject's philosophy of life as

6 A specimen analysis will be found in the Appendix, p. 102.
it was applied by the subject herself to those aspects of living that were most significant to her at the time.

**TABLE XXI**

**DIFFERENCES BETWEEN PAIRED ELEMENTARY GROUPS ON EIGHT SIGNIFICANT ITEMS OF THE MTAI**

*N = 25 pairs*

<table>
<thead>
<tr>
<th>Items of MTAI</th>
<th>Answers</th>
<th>Frequencies</th>
<th>Chi square</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>High</td>
<td>Low</td>
<td></td>
</tr>
<tr>
<td>5. Teaching never gets monotonous.</td>
<td>SA,A</td>
<td>20</td>
<td>10</td>
<td>6.750</td>
</tr>
<tr>
<td></td>
<td>D,SD</td>
<td>5</td>
<td>15</td>
<td></td>
</tr>
<tr>
<td>101. Most pupils are considerate of their teachers.</td>
<td>*SA,A</td>
<td>23</td>
<td>14</td>
<td>6.752</td>
</tr>
<tr>
<td></td>
<td>*U,D,SD</td>
<td>2</td>
<td>11</td>
<td></td>
</tr>
<tr>
<td>39. To maintain good discipline in a classroom a teacher needs to be &quot;hard-boiled.&quot;</td>
<td>SD</td>
<td>19</td>
<td>9</td>
<td>6.575</td>
</tr>
<tr>
<td></td>
<td>D</td>
<td>6</td>
<td>16</td>
<td></td>
</tr>
<tr>
<td>38. The majority of children take their responsibilities seriously.</td>
<td>SA,A</td>
<td>19</td>
<td>10</td>
<td>5.254</td>
</tr>
<tr>
<td></td>
<td>U,D</td>
<td>6</td>
<td>15</td>
<td></td>
</tr>
<tr>
<td>54. Most children lack common courtesy toward adults.</td>
<td>SA,A</td>
<td>7</td>
<td>16</td>
<td>5.153</td>
</tr>
<tr>
<td></td>
<td>U,D,SD</td>
<td>18</td>
<td>9</td>
<td></td>
</tr>
<tr>
<td>89. Teachers who are liked best probably have a better understanding of their pupils.</td>
<td>SA</td>
<td>11</td>
<td>3</td>
<td>4.861</td>
</tr>
<tr>
<td></td>
<td>*A,U,D</td>
<td>14</td>
<td>22</td>
<td></td>
</tr>
<tr>
<td>145. Young people today are just as good as those of the past generation.</td>
<td>*SA</td>
<td>16</td>
<td>8</td>
<td>4.083</td>
</tr>
<tr>
<td></td>
<td>A</td>
<td>8</td>
<td>16</td>
<td></td>
</tr>
<tr>
<td>90. Most pupils try to make things easier for their teachers.</td>
<td>SA,A</td>
<td>16</td>
<td>8</td>
<td>3.926</td>
</tr>
<tr>
<td></td>
<td>U,D,SD</td>
<td>9</td>
<td>17</td>
<td></td>
</tr>
</tbody>
</table>

SA-strongly agree; A-agree; U-undecided; D-disagree; SD-strongly disagree.
*Grouping of answers agrees with that of the authors of the inventory.
Yates correction was used in the computation of chi square.

Ten pairs of teachers were selected at random from the secondary group.

The plan was to set up a "scoring key" empirically on the basis of the
statements of these ten high-rated and ten low-rated teachers--a total of 120 statements for each group. Careful inspection revealed a certain con-
structive attitude running through the sequence analyses of the high group that was not evidenced in the other. Though not very clear at first, its presence aided the investigator in assembling the types of statements that appeared predominantly in each group. Statements that appeared only occasion­ally in either group and expressed a basic disposition contrary to the predominant one, were transferred to the opposite set. Thus there were gradually built up two opposing classes of statements, one of which repre­sented most of the dispositions expressed by the high group, and the other, those expressed by the low group. The former were designated "plus"; the latter, "minus."

It was found that there was a strong contrast in mental set between the two groups, which had to do with the way they saw reality and the de­mands it made on them. Further inspection showed that the aspects of life toward which this basic positive or negative attitude was expressed could be classified in a limited number of categories. For example, the teachers were concerned with achievement, and with rare exceptions the import of their stories revealed their basic outlook on life with reference both to success and to failure. Another aspect of life to which they reverted re­peatedly was loss in its various forms: disappointments, injustice, hard­ship. There was involvement with life as a whole, its values, the right and wrong of things, and finally, with other people. For convenience of scoring the following five categories were, therefore, set up: success, failure, loss, life and its obligation, other people. Each of the assembled
"plus" and "minus" statements was assigned to the category to which it had reference. Thus the basic positive or negative mental set of the subjects, as expressed through their stories, was defined as it applied to the more significant areas of the subjects' lives.

After these categories of statements had been tentatively set up with the secondary teachers, a trial test was made to determine whether they could be used with the elementary teachers to distinguish the groups according to their rating. Twenty cases were selected at random from among the latter. It was not known which of these cases were high and which low, nor were pairs identified. When the experimenter tried to predict blindly, on the basis of the categorized statements alone, which were rated high and which low, the attempt was only partially successful. While the same positive or negative attitude was apparent in the second group of TAT's, it was not sufficiently defined by the prepared key to make its recognition possible in every statement of a given case.

This second group was, therefore, examined for differentiating characteristics in the same way that the first had been, and the results were combined with the original. Thus, the final "scoring standard," as this pattern of categorized statements will henceforth be designated, was set up on the basis of forty cases, twenty secondary and twenty elementary. Table XXII shows illustrative contrasting statements in two of the categories.7 The phraseology follows closely that of the sequence analysis.

7The complete scoring standard will be presented and discussed in Chapter IV.
### Table XXII

**ILLUSTRATIVE STATEMENTS FROM TWO CATEGORIES OF THE SCORING STANDARD**

<table>
<thead>
<tr>
<th>Category A - Success</th>
<th>Category D - Life and its obligations</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Plus</strong></td>
<td><strong>Minus</strong></td>
</tr>
<tr>
<td>The goal is attainable in spite of difficulties; one must go after it.</td>
<td>The goal is doubtful of achievement; it is merely hoped for.</td>
</tr>
<tr>
<td>Achievement depends on work, sacrifice, planning.</td>
<td>Success just comes; it comes easily or in some unrealistic way.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><strong>Plus</strong></th>
<th><strong>Minus</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>Duties are undertaken and decisions made for rational causes; they are the right thing to do; they help others.</td>
<td>Duties are undertaken and decisions made because there is no way out, or just to please.</td>
</tr>
<tr>
<td>Right conduct is rewarded; love is a help in doing right.</td>
<td>Right conduct brings failure, loss; love makes life easier.</td>
</tr>
</tbody>
</table>

The remaining sixty cases were then evaluated according to this scoring standard. The cases were not identified as to pairs nor rating. Each statement in each subject's sequence analysis was marked either **plus** or **minus**, according to the scoring standard, and, on the basis of the total number of positive statements, each case was assigned to the high or the low group. Only two cases were not correctly assigned in this attempt at prediction. Examination showed that these two failures were due to faults in the original sequential analysis, and when these faults were corrected, the two cases plainly fell into the classes in which they belonged.

These results demonstrated that the **plus** and **minus** statements
represented discriminating differences between the high- and the low-rated teachers. As a further test of the results, the TAT protocols and the scoring standards were given to two other judges, who scored the protocols independently. Prediction was 100% correct; that is, both judges accurately placed every teacher in the high or low group on the basis of her plus and minus scores. On the 1200 stories taken as a whole, there was 97.2% and 96.6% agreement between the investigator and judges A and B respectively. Agreement between judges A and B was 94.3%. (Table XXIII) In no case was there sufficient disagreement with regard to any one subject to cause that teacher to be incorrectly placed in the high or the low group.

**TABLE XXIII**

**SCORING RELIABILITY OF THE TAT**

<table>
<thead>
<tr>
<th>Judges</th>
<th>Agreement on scoring 1200 TAT stories</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Number of differences</td>
</tr>
<tr>
<td>A &amp; investigator</td>
<td>34</td>
</tr>
<tr>
<td>B &amp; investigator</td>
<td>41</td>
</tr>
<tr>
<td>A &amp; B</td>
<td>68</td>
</tr>
</tbody>
</table>

In an attempt to arrive at a statistical figure to indicate the significance of the differences between the high and the low teachers, the Sign Test was applied to the pairs of scores. Since all the differences were in the same direction, that is, in every pair of teachers the high-rated one received more plus scores, $x$ (the number of fewer signs) was 0. For $N = 25$, an $x$ of this size is significant beyond the .1% level of confidence.
The results were the same in both grade and high school groups. In Table XXIV are presented the positive scores made by the 25 pairs of elementary and 25 pairs of secondary teachers.

### Table XXIV

**Scores of high and low groups on the TAT**

<table>
<thead>
<tr>
<th>Pairs</th>
<th>Positive scores Elementary teachers</th>
<th>d</th>
<th>Positive scores Secondary teachers</th>
<th>d</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>High-rated</td>
<td>Low-rated</td>
<td></td>
<td>High-rated</td>
</tr>
<tr>
<td>1</td>
<td>12</td>
<td>1</td>
<td>11</td>
<td>11</td>
</tr>
<tr>
<td>2</td>
<td>9</td>
<td>3</td>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td>3</td>
<td>12</td>
<td>1</td>
<td>11</td>
<td>12</td>
</tr>
<tr>
<td>4</td>
<td>10</td>
<td>0</td>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td>5</td>
<td>8</td>
<td>3</td>
<td>9</td>
<td>9</td>
</tr>
<tr>
<td>6</td>
<td>9</td>
<td>3</td>
<td>6</td>
<td>9</td>
</tr>
<tr>
<td>7</td>
<td>12</td>
<td>2</td>
<td>10</td>
<td>12</td>
</tr>
<tr>
<td>8</td>
<td>10</td>
<td>2</td>
<td>8</td>
<td>10</td>
</tr>
<tr>
<td>9</td>
<td>10</td>
<td>4</td>
<td>6</td>
<td>12</td>
</tr>
<tr>
<td>10</td>
<td>11</td>
<td>1</td>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td>11</td>
<td>9</td>
<td>2</td>
<td>7</td>
<td>12</td>
</tr>
<tr>
<td>12</td>
<td>9</td>
<td>3</td>
<td>6</td>
<td>11</td>
</tr>
<tr>
<td>13</td>
<td>11</td>
<td>1</td>
<td>10</td>
<td>12</td>
</tr>
<tr>
<td>14</td>
<td>8</td>
<td>2</td>
<td>6</td>
<td>10</td>
</tr>
<tr>
<td>15</td>
<td>10</td>
<td>1</td>
<td>9</td>
<td>9</td>
</tr>
<tr>
<td>16</td>
<td>11</td>
<td>0</td>
<td>11</td>
<td>9</td>
</tr>
<tr>
<td>17</td>
<td>8</td>
<td>4</td>
<td>4</td>
<td>9</td>
</tr>
<tr>
<td>18</td>
<td>12</td>
<td>2</td>
<td>10</td>
<td>9</td>
</tr>
<tr>
<td>19</td>
<td>10</td>
<td>0</td>
<td>10</td>
<td>12</td>
</tr>
<tr>
<td>20</td>
<td>11</td>
<td>1</td>
<td>10</td>
<td>9</td>
</tr>
<tr>
<td>21</td>
<td>9</td>
<td>3</td>
<td>6</td>
<td>11</td>
</tr>
<tr>
<td>22</td>
<td>12</td>
<td>3</td>
<td>9</td>
<td>12</td>
</tr>
<tr>
<td>23</td>
<td>10</td>
<td>3</td>
<td>7</td>
<td>12</td>
</tr>
<tr>
<td>24</td>
<td>11</td>
<td>0</td>
<td>11</td>
<td>10</td>
</tr>
<tr>
<td>25</td>
<td>11</td>
<td>3</td>
<td>8</td>
<td>9</td>
</tr>
</tbody>
</table>

x (number of minus signs in d) 0
Level of significance .1%
To determine which questions of the rating scale were most significant in the selection of the teachers possessing the positive characteristics, phi coefficients were computed to find the relationships between the teachers' scores on the TAT and their rating on each of the seven questions. The two categories of TAT scores were the number of teachers making scores of 0-4 and the number making scores of 8-12. (There were no scores between 4 and 8.) The ratings were dichotomized at the 50th centile. Table XXV shows these correlations for the elementary and the secondary teachers. The coefficients ranged from .37 to .96.\(^8\) Lowest relationships were with Questions IV and VI on both levels; highest, with Questions I and VII in the grade school and with Question I in the high school.

**TABLE XXV**

**RELATIONSHIPS (PHI COEFFICIENTS) BETWEEN TAT SCORES AND RATINGS ON THE SEVEN QUESTIONS OF THE SCALE**

<table>
<thead>
<tr>
<th>N</th>
<th>Level</th>
<th>Questions of rating scale</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>I</td>
</tr>
<tr>
<td>50</td>
<td>Elementary</td>
<td>.96</td>
</tr>
<tr>
<td>50</td>
<td>Secondary</td>
<td>.96</td>
</tr>
</tbody>
</table>

All coefficients significant beyond 1% level.

\(^8\)^Phi coefficients are restricted in size and are, therefore, lower than Pearson r's. The phi coefficients in this study are further reduced by the use of Yates’ correction for continuity. These coefficients may, accordingly, be looked upon as conservative estimates of the correlation between the given variables.
Question I, "How well do you like your teacher," was apparently the real criterion in this study for both the elementary and the secondary teachers.

It could safely be concluded from these results that the contrasting statements in the scoring standard represented true differences between the teachers as they were rated by their pupils. Since the object of the study was to discover such differentiating characteristics and to understand their nature, Chapter IV will be devoted to a more thorough analysis of them.
CHAPTER IV

INTERPRETATION OF THE TAT FINDINGS

It was suggested at the beginning of this study that the characteristic differences between "good" and "poor" teachers cannot be discovered by behavior sampling techniques. This research has demonstrated that there are differences between teachers that are not only significant, but clearly discriminating, and that these differences can be found by means of TAT Sequential Analysis.

What are these differences, and what part do they play in the teacher's personality? Why did the pupils so consistently select as "good" or "poor" teachers those characterized by the qualities revealed in this study? These questions will be discussed in this chapter in an attempt to arrive at a clearer understanding of the characteristics indicated by the TAT and their effect on the relationship between the teacher and her pupils.

To make the findings of this study more intelligible, it is necessary to explore somewhat thoroughly two distinct avenues of thought. These are best expressed by the following questions: (1) What is the nature, in general, of results derived through TAT Sequential Analysis? (2) What is the relationship of the specific results found in this study to the pupils' perception and evaluation of their teachers?

Before approaching a discussion of either of these questions, it will be helpful to examine the detailed descriptive account of the factual
findings of the investigation. These findings are given in Tables XXVI to XXX. The tables constitute the scoring standards as finally set up and according to which the protocols of the high and low groups were scored and distinguished.

As previously explained, the meaning of each story was set down in a statement; and, according to their basic pervading import the statements were assembled in two contrasting groups, which were designated as plus and minus respectively. The groups were further subdivided into five categories, which named the predominant areas of life toward which the basic positive or negative attitude was expressed. These categories are: success, failure, loss, life and its obligations, other people. The categories thus formed are somewhat arbitrary; but they are nonetheless dependent upon the data given in the study. They were established after the sequence analysis had been completed and on the basis of this analysis. They are, therefore, empirical in nature. The statements were designated as positive or negative according to whether they were of the type that appeared predominantly in the sequence analyses of the high- or the low-rated group of teachers. This predominance was ascertained by inspection. It was verified in the predictive part of the study and in the scores of the high and low groups.
### TABLE XXVI

**HABITUAL BASIC DISPOSITION AS EXPRESSED TOWARD SUCCESS**

**Category A.** This category is used when the story involves the achievement of something; success is aimed at or attained; it includes overcoming temptations and fears, working through a problem.

<table>
<thead>
<tr>
<th><strong>Scoring standard</strong></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Plus</strong></td>
<td><strong>Minus</strong></td>
</tr>
<tr>
<td>The goal is clearly seen; it is attainable in spite of difficulties; one must go after it.</td>
<td>The goal is not clearly seen; there is uncertainty, doubt as to what course to take; fear to enter upon the course; compromise between two courses, keeping both; putting off the solution of a problem.</td>
</tr>
<tr>
<td>Achievement depends on constructive action on the achiever's part: effort, work, bearing the pain and weariness work entails, determination, persistence, overcoming obstacles, sacrifice, facing danger or risk, preparation, planning, prudence, doing the right thing, interest in worthwhile things, using opportunities. One receives advice, help, encouragement or inspiration from others. One works with others. (Wherever others are involved, effort on one's own part must be explicitly stated or clearly implied.) Self is overcome because it is the right thing to do; or for supernatural reasons; or for love of someone; or with supernatural help. (Difficulty is recognized.)</td>
<td>The goal is seen as foolish, or wrong, or unrealistic; or it is doubtful of achievement; or it is merely accepted because of circumstances. The goal is relinquished because of pain, danger; a higher goal is relinquished in favor of a lesser one; one can be satisfied with an easy or ordinary goal. A goal is merely hoped for, a dream. A goal is achieved, but there is no explicit constructive action on the achiever's part; success just comes; it comes easily or in some unrealistic way; it depends on others; one puts one's whole confidence in them; it comes through prayer with no effort; it comes miraculously; it comes with worry or anxiety; it comes in the form of reward at the end of a happy life. Resolutions are made on the spur of the moment, without foresight or planning.</td>
</tr>
<tr>
<td>Resolutions to achieve require a recognition of difficulties involved; they require deliberation or planning.</td>
<td></td>
</tr>
<tr>
<td>Scoring standard</td>
<td></td>
</tr>
<tr>
<td>------------------</td>
<td>---</td>
</tr>
<tr>
<td><strong>Plus</strong></td>
<td><strong>Minus</strong></td>
</tr>
<tr>
<td>Failure is the result of inadequacy in oneself.</td>
<td>Failure just happens, with or without explicit effort.</td>
</tr>
<tr>
<td>Failure can be overcome by constructive action: any of the means listed in Category A; or, with another chance; with supernatural help; by changing the goal to one more realistic; through learning by mistakes; by drawing good from evil; for the sake of a loved one.</td>
<td>Further failure is expected.</td>
</tr>
<tr>
<td>After failure one can still be happy if one knows that one has done the right thing. (not just tried)</td>
<td>Failure is due to others or to anything but oneself.</td>
</tr>
<tr>
<td></td>
<td>Failure cannot be overcome or is not overcome: one makes the best of it; is resigned, does nothing; tries to be happy in spite of failure; quickly forgets it; assures herself that she has tried; just hopes; is deserted, tearful, worried; seeks to escape, actually or in dreams; continues to fail in the eyes of others even when she does better; others comfort her.</td>
</tr>
<tr>
<td></td>
<td>Failure is not admitted: there is no need to fear; everything will turn out well in spite of failure.</td>
</tr>
<tr>
<td></td>
<td>Failure is overcome with no constructive action: one is rescued; one's problem is solved by someone else; one simply corrects or is willing to correct one's mistakes; one succeeds eventually; one does better but with little success; one returns to duties as a result of others' prayers.</td>
</tr>
<tr>
<td></td>
<td>What was taken for success proves to be hollow or unsatisfying.</td>
</tr>
</tbody>
</table>
TABLE XXVIII
HABITUAL BASIC DISPOSITION AS EXPRESSED TOWARD LOSS

Category C. This category is used when the story involves a giving up that hurts; any loss, setback, disappointment, separation, deprivation, sorrow, hardship which is not failure to achieve; it is the effect of God's Will, circumstances, or the actions of others.

<table>
<thead>
<tr>
<th>Scoring standard</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Plus</strong></td>
<td><strong>Minus</strong></td>
</tr>
<tr>
<td>Something is actively done about the loss: one gives up of her own accord, even though reluctantly; one actively accepts the loss.</td>
<td>Loss is due to others; it is reason for rebellion, resentment, or other negative emotion: worry, fears, tears.</td>
</tr>
<tr>
<td>This activity is shown by: overcoming negative emotion; substituting work or doing good; seeing good come out of the loss; seeing the possibility of evil in the object lost.</td>
<td>Loss is not accepted: one never gets over it; it always remains a tragedy; she substitutes something pleasurable; escapes in dreams; is hardened, inured to her loss; one always seeks or compromises.</td>
</tr>
<tr>
<td>Loss is accepted as required: by the nature of things; or for the good of others or of oneself; or by one's duty.</td>
<td>Loss is accepted passively: one can't get out of it; is resigned after a while; just loses, is deprived; it just happens; one prays but there is no positive effect; one tries to be happy with what one has left but still goes on seeking the lost object; one is glad when the trouble is over, without action on her part; gets over it easily or with no constructive action.</td>
</tr>
<tr>
<td>There is hope in spite of the loss; prayer can save one from loss.</td>
<td>One is content to profit by another's loss.</td>
</tr>
</tbody>
</table>

(When something is sacrificed in order to achieve or something is undertaken to overcome a loss, the story enters the success or failure category.)
### TABLE XXIX

**HABITUAL BASIC DISPOSITION**

**AS EXPRESSED TOWARD LIFE AND ITS OBLIGATIONS**

**Category D.** This category is used when the story does not belong to any of the previous categories, but indicates the writer's attitude toward life, values, and duties; it shows an attitude toward work, rest, or reward without relating to achievement.

<table>
<thead>
<tr>
<th>Scoring standard</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Plus</strong></td>
</tr>
<tr>
<td>Life is seen as it actually is; the statement is objectively true; there is hope in God and in prayer; first things come first; God is loved or worshiped for Himself, because He is good; logical reasons are given for actions and for emotional reactions.</td>
</tr>
<tr>
<td>Work is tiring, but this is accepted; it is not overwhelming; there is time for rest; work goes on; it brings reward.</td>
</tr>
<tr>
<td>Duties are undertaken and decisions are made for rational causes: they are the right thing to do; they help others.</td>
</tr>
<tr>
<td>Right conduct is rewarded; love is a help in doing right.</td>
</tr>
<tr>
<td>Wrongdoing is punished; wrongdoing brings harm to self or to others; wrongdoing is regretted because it is wrong.</td>
</tr>
<tr>
<td>Duties are ignored, neglected, escaped; fun or freedom is preferred. Right conduct brings failure, loss. Love makes life easier (no effort). Wrong is condoned, protected; wrong is regretted for any reason other than conviction of its wrongness; or no reason is given and no effort made to make up for it.</td>
</tr>
</tbody>
</table>
TABLE XXX
HABITUAL BASIC DISPOSITION
AS EXPRESSED TOWARD OTHER PEOPLE

Category E. This category is used when the story shows primarily an attitude toward others and does not belong in any previous category; usually simple statements (in the sequence analysis) of what others do, should do, or of what she should do to them.

<table>
<thead>
<tr>
<th>Scoring standard</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Plus</strong></td>
</tr>
<tr>
<td>Others are good, helpful; others are forgiving and trustful when one makes a constructive effort; difficulties with others can be cleared up.</td>
</tr>
<tr>
<td>One is edified by others when they do the right thing; one is true to friends.</td>
</tr>
<tr>
<td>Others should be helped, repaid, rewarded; one should sacrifice for others.</td>
</tr>
<tr>
<td>One learns from the mistakes of others; one profits by their advice.</td>
</tr>
<tr>
<td>As a teacher, one helps the young; saves, encourages, and supports them; the young respond to one's efforts; they have high ideals and are willing to face difficulties in attaining them.</td>
</tr>
</tbody>
</table>

(Achievement with or through others is in the success category; failure with or through others is in the failure category; injustice or loss through others is in the loss category; doing right or wrong with or because of others is in the obligations category.)
The Nature of TAT Findings

Examination of the above tables reveals that they do not describe external behavior. Each statement is an expression of a way of thinking or of seeing the world, a basic frame of mind. From the contrast between the plus and minus statements, it is apparent also that people differ in their ways of looking at things. A person may have a basic frame of reference which causes him to say, "A goal can be achieved in spite of difficulties; achievement depends on work and planning." Or, he may look on striving after a goal as "foolish and unrealistic; success just comes." Again, one person may have an orientation toward others which is basically, "Others are good, helpful; they forgive when one makes a constructive effort"; while another may be disposed to think, "Others are frustrating, interfering; they do not overlook one's defects even when one does right."

It is further apparent from the tables that the characteristics listed under plus in every category are such as would be objectively selected as "desirable" ways of looking at things. Yet the low-rated teachers in the present study did not reveal this desirable way of seeing their world. The attitude of this group is given under minus, which, in every category, expresses a view of the world that they would be expected rationally to repudiate. The teachers were, of course, not aware of what they were revealing in their TAT stories, and it is safe to say that in most cases they were not fully aware of the given attitude in themselves.

Still, they were not totally unconscious of it. To understand this, one must know what sequential analysis is, and how the subject reacts to
it in a clinical situation. In the first place, the analyst by this method does not impute to the subject any attitudes or characteristics. The subject speaks for himself and the analyst merely abstracts the import of what the person says. The subject may not realize at the moment of writing what he is asserting by his story. However, when the meaning of his story is read back to him later, he can readily recognize that that is what his story actually says; and a little introspection makes it possible for him to see also that it is his usual way of looking at things, though he may seldom, if ever, have adverted to the fact. In other words, the import of the story, while given unconsciously at the time of writing, is not a mystery to the writer, and if he possesses any degree of intellectual honesty, he does not deny it when confronted with it.

Consideration of the fact that the positive-negative dichotomy of the tables was determined by the pupils' ratings strongly suggests also that the results of TAT Sequential Analysis, while not in the form of overt actions, are really the antecedents of action. These statements represent the only apparent distinction between the teachers who were rated at the two extremes. Do these diverse ways of seeing the world result in behavior that accounts for the differences in rating? It is reasonable to expect a person's actions to be consistent with his way of thinking. This line of reasoning leads directly to the consideration of the second question proposed at the beginning of this chapter, "What is the relationship between the specific results found in this study and the pupils' perception and evaluation of their teachers?"
Teacher Characteristics as Evaluated by the Pupils

The pupils, of course, did not rate their teachers on the characteristics revealed by the TAT. They observed whether or not their teacher taught the lessons in an interesting way, explained things clearly, was kind and understanding as she did so, kept order and managed everything well so that they came to like her and the lessons. These are all externals; and by rating her on these they discriminated between people who had clearly different interior dispositions with regard to life in general. In other words, the predictive data and the criterion are in two entirely different dimensions.

Examination of the categories involved in Tables XXVI to XXX reveals that they represent very important and fundamental aspects of life: success and failure, possession and loss, the conflict between duty and pleasure, relationships with reality and with other people. These are contingencies that permeate all of life's situations, just as they permeated the situations which the teachers created in their stories. And the frame of mind with which each one set up and solved the problems in her stories must perforce be the same disposition with which she handles the problems and circumstances of real life. Just as each teacher approached the story-writing task with a definite mind-set and internal convictions, so too she approaches the ordinary tasks of the day.

When one realizes this fact and examines the qualities on which the teachers were rated by the pupils, the connection becomes apparent in many respects. A teacher, for example, may be rated low on "explaining the
lessons." It is not difficult to see how one who is not really convinced that it takes preparation and work to succeed will not put forth the effort that is required to make things clear to immature minds. Nor is it hard to understand that a person who sees others as unfriendly and frustrating will tend to look upon all the actions of children as threats to her security and strike out to protect herself; and her pupils will, consequently, find her "low" in understanding. When duties are undertaken because there is no way out, and work is looked upon as drudgery, lessons will be poorly prepared and boring, and there will be no enthusiasm to catch fire in young hearts. The teacher who is absorbed by her anxieties or resentful over losses will hardly be the one to approach her classroom problems pleasantly and be rated "high" on her cheerful management of her class.

Neither is it difficult to understand why the person who sees her goal clearly, and knows that it cannot be reached without hard work and sacrifice, will begin anew without frustration when her well-laid plans do not succeed immediately. Her pupils find her painstakingly repeating her explanations, going out of her way to help them in their difficulties. The teacher who sees the positive side of her losses and disappointments is likely to carry over that attitude into her classroom; and her pupils rate her "high" on cheerfulness and the pleasant way in which she keeps order. She who looks on others as cooperative and responsive will undoubtedly demonstrate that disposition toward her pupils and win their esteem and affection.

Not all the implications of the differences uncovered in the study are so readily apparent, however. Sometimes the distinction between the plus and minus qualities is exceedingly fine, or, at least, it may seem so when
an attempt is made to apply a given statement to concrete instances. Such is the case, for example, in the reasons for loving and worshiping God, as they appear in Category D. The plus statement is: "God is worshiped for Himself, because He is good"; the minus: "God is worshiped because of loveliness, sweetness, beauty." The discrimination here lies in the difference between concentrating on God and concentrating on self, which is a tremendous distinction in itself, but one that can easily escape notice in a test analysis. When it is seen for what it really is, its consequences upon the teacher in her classroom can be more readily understood.

Another very delicate distinction exists between loving others for what they are in themselves and loving them for what they give. There is also need to distinguish precisely between a healthy dependence on prayer and on people--which is not blind to the necessity of work and sacrifice on one's own part--and the passively dependent attitude which looks for help and success without a sane recognition of the demands that inhere in achievement.

This ability or inability to see and to comprehend the inherent demands of reality seems to be the very core of the contrast between the high and the low teachers. Throughout the TAT's of the high group there is a steady, persistent and, at the same time, rationally approved and optimistically active, way of responding to the world as it is. This is evidenced in statements such as the following, appearing in the protocols of the high group:

One must go after a goal; achievement depends on constructive action, overcoming obstacles, using opportunities. Failure can be overcome by changing the goal to one more realistic. Something
can be done about a loss; good can come out of it. First things come first. There are logical reasons for actions and emotional reactions. Wrongdoing is regretted because it is wrong; duties are undertaken because they are the right thing to do. Others are helpful.

In the opposite picture there is a lack of realism and a dearth of constructive action:

There is uncertainty, fear to enter upon a course. The goal is foolish, unrealistic; it is only a dream. Success or failure just happens. Loss always remains a tragedy; one can't get out of it. Everything always turns out all right. Work is drudgery and one just gets tired. Experiences that hurt are to be avoided. Reward comes without work. Others are the sources of one's troubles.

The teacher who is rated high by her pupils sees the world as it is: stern, demanding, and at times painful and disappointing. But at the same time it is a hopeful world, filled with worthwhile things to do and to achieve, and with people who are helpful and encouraging. In order to attain success or happiness certain things need to be done and one goes ahead with them, foreseeing the difficulties and planning how to overcome them. Others are helpful at times, but success depends primarily on one's own prudence and action.

It is a different world to the teachers at the opposite extreme. Life for them is somehow all wrong; and there is no reason for its being so, nor can aught be done about it. Or, life is fantastically devoid of anything that demands the use of foresight or the exertion of persistent energy. In either case there is a passive acceptance or a passive, though deep-seated, resentment. What is always lacking is the recognition of the need for wholehearted, constructive action.

Furthermore, all things are perceived by these teachers from the
viewpoint of their own interest; and this is true whether they speak of
their relationships to God, the world and other people, or of the immediate
circumstances of their lives. God is worshiped for what He gives, rather
than for what He is; duties are undertaken because they are pleasurable or
convenient, not because they are right; people are good for what they con-
tribute, not for what they are.

These ways of seeing things are not constructive because they do not
correspond to the objective truth. The high-rated teachers seem to discern
clearly what a situation demands, and this view apparently dictates action
consonant with the scheme of things. Seeing the world as it really is leads
to logical judgments about it, followed by the necessary action. That a
realistic frame of reference carries over into action must be concluded
from the rating that these teachers received from their pupils, who rated
them on what they did.

It is somewhat difficult to analyze just what happens in the case of
the low-rated teachers. Do they not know and understand the truth in the
same way as the others? There are several possibilities. They may see
what a situation demands, but rebel against it; they may fail to recognize
it because of emotional or defensive habits that have been built up; or
they may simply not comprehend the reality because of lack of experience.
The first or the second possibility may be present where there is doubt,
anxiety, resentment, or irrational optimism. The third may account for the
incomprehensible naiveté with which solutions are sometimes proffered. In
all cases these teachers seem to have developed habits of closing their
minds to both the limitations and the opportunities of life as it is, or of
chafing under the limitations that they see and of refusing to command the
called-for action. This way of reacting to life, as it appears in the TAT
stories, is likewise carried over to the classroom, at least to an extent
sufficiently great to cause the pupils of these teachers to rate them low
on what they do.

The explanation of this is to be found in the nature of the function
of habit. As was pointed out before, the teachers did not advert to this
inner disposition as they wrote their stories. The attitude that shone
through the story was habitual; it operated easily without conscious adver-
tence. In the same way, that habitual way of seeing things directs all their
actions, very frequently, no doubt, without clear awareness on their part.
But the pupils who are with them day by day, hour after hour, observe their
actions and sense the attitude behind them. Though they may be unable to
give a correct explanation for their feelings and their likings, they react
to the teacher's personality as it really is. That fact is demonstrated in
this study by the clear distinction that appeared in the personalities that
were rated by them at opposite poles, and particularly by the high correla-
tion between their "liking the teacher" and the TAT scores.

This truth, that the teacher will act according to her habitual mind-
set as a general rule, though she may depart from the usual pattern for
short periods when an adult observer is in the classroom, may account for
the greater reliabilities that have generally been found in the ratings of
pupils. Their rating is based, not on a sample of classroom behavior, but
on all of it; and it is probably, therefore, a truer estimate of a teach-
er's personality.
Thus, a fundamentally realistic and constructive disposition, as defined in the scoring standard supported by the predictive and statistical procedures of this study, would seem to be the determining characteristic of teachers rated high by their pupils. An objective view is apparently the prerequisite to constructive and persevering action, and to the setting up and maintenance of adequate relationships with people and with the contingencies of life as the teacher meets them in her school situation. The almost perfect correlation that exists between the possession of this attribute and the pupils' evaluation argues that it is the necessary characteristic of a good teacher as her pupils see her.

It would be a mistake to think, on the ground of the foregoing, however, that the teacher who is rated high by her pupils is a paragon of virtue. Such is not actually the case. It is conceivable that the realistic and persistent action that characterizes the high-rated teacher may be diverted to unworthy causes, so that there may be some ways in which the teacher may not be rated high as a person. She may at any point deliberately act in a manner that is not consonant with her objective view; or she may fail to see clearly in some respects because of emotional involvement. This may account for the conflict that sometimes appears in the TAT of a high-rated teacher. Her conflict may be caused by her striving after a goal that is incompatible with some aspect of her life, but not necessarily with her goals in the classroom. Here she apparently sees objectively what the situation demands and acts in accordance with that view, to an extent, at least, to merit a high rating by her pupils.

Similarly, the teacher who is rated low is not necessarily inadequate
in every respect. A lack of realism in one's view of the world and the resultant lack of constructive action, to the degree described here, may not cause failure in all aspects of life, though they may be conducive to serious difficulties in one's relationships generally. The psychological requirements of a good teacher, as her pupils see her, are of an exceedingly high order, and it is to be expected that many will not attain to them, though they may lead a life that is not only morally irreproachable, but even productive of good.

The characteristics, as they are defined in this research, are based on an empirical study of the teachers. Conclusions as to whether these qualities are also the requisites of a good moral or religious life, or of conduct desirable in other specific circumstances, must be determined by research that is based on the appropriate criterion.
CHAPTER V

DISCUSSION

Certain broad conclusions can be drawn from the results of this investigation. The first is that a religious teacher's habitual basic disposition toward life in general is a greater factor in her acceptance by her pupils than are the truths and attitudes to which she gives conscious expression. These may be in agreement; but when they are not, it is the former that prevails. This is apparent from the much more conclusive results that were obtained with the TAT than with the MTAI.

The second conclusion is that TAT Sequential Analysis is a powerful technique for indicating the underlying basic disposition that influences a person's habitual reactions to the contingencies of life. This conclusion becomes evident when one compares the findings of the TAT with those of other instruments. In this study, for instance, there were certain significant differences revealed by the MTAI with some classes of teachers. Yet there was much overlapping of attitudes even within these classes. Some of the teachers who were rated highest by their pupils made the lowest scores on the MTAI and vice versa. This was not true of the TAT results, in which there was not only no overlapping of scores, but the scores of the criterion groups were actually at the extremes.

Another conclusion is that pupils, given conditions that allow an honest expression of opinion, will, as a group, rate their teachers in an
amazingly astute and consistent way. In the present sample of one hundred teachers, there was not a single discrepancy between the TAT scoring and the pupils' ratings. This was true in the case of fifth grade pupils as well as of the twelfth.

It can further be concluded that the habitual basic disposition that is required to produce a high-rated teacher is essentially the same in the elementary and the secondary school. This judgment holds for all teachers, furthermore, whether they score high or low in intelligence.

The latter consideration leads to a fifth conclusion, namely, that, while intelligence is an important factor, at least in the way high school pupils rate their teachers, it is secondary to the personal qualities here revealed. It is interesting to note, in this connection, that the high school teacher's ability to explain clearly not only was most closely associated with intelligence, but also correlated highly, second only to "being liked," with the personality characteristics.

The latter fact supports the conjecture that the habitual basic disposition of the teacher, as revealed by the TAT, is more than just a factor in her acceptance by her pupils. It gives reasonable assurance about the teacher's actual performance, and thus can be regarded as a fairly accurate indicator of teaching efficiency.

There are certain limitations that need to be pointed out with regard to the conclusions of the present investigation. While the study revealed clearly discriminating differences between the high- and the low-rated teachers, it does not follow necessarily that these differences as here defined provide an infallible predictive measure of a teacher's evaluation by
her pupils. This investigation was carried out with extreme groups, a procedure which was calculated to maximize the differences so that they could be more readily identified. Further research is needed with the two hundred cases who were rated between the extremes, before it can be known to what degree the results of the TAT can predict a teacher's rating. The fact that those who were rated at the upper and lower ends also scored at the extremes on the TAT suggests that the middle group might present a range of scores falling between these extremes. Actual experimentation with this group is required to support this hypothesis.

Another limitation, which is not peculiar to this research but inheres in all personality studies, is suggested by the question of the permanence or consistency of these characteristics within a teacher. While a disposition appears to be habitual and basic to the personality at the present moment, will it continue to be so? It may be firmly rooted in some; possibly less so in others. An answer to this question would require further experimentation with the same group of teachers at various times and under differing circumstances.

There is still a third limitation which arises in conjunction with one of the problems that was proposed at the beginning of this investigation. It had been hoped that the results would bear some implications for the education of the student teachers. Whether and to what degree the characteristics found are of such a nature that their development can be promoted by education, particularly the education received after the age of admission to a teacher-education institution, is a question that remains unanswered. The teachers in the present study, especially those in the paired samples,
were fairly homogeneous in education and training. Do the differences re-
vealed in the study exist before their entrance into the institution? If
so, to what extent can the right disposition be developed by direct effort
in their training? These questions require research of a longitudinal na-
ture, and until they are answered, only very tentative inferences can be
drawn from the present investigation for the education of young teachers.

The MTAI indicated that at least on some grade levels the teachers' atti
dudes toward children and teaching are important factors. The improve-
ment of these attitudes could possibly be effected by better understanding
of the nature both of children and of teaching, as well as by better super-
vised and guided initial classroom experiences. Also a thorough knowledge
of the subject matter to be taught would undoubtedly increase the young
teacher's feeling of security and her chances of making a successful and
happy beginning in her profession. Consequently, improvement along both of
these lines in the teacher's education would be desirable.

However, these factors alone will not develop the basic disposition
that has been shown here to distinguish the high-rated teachers. It may
be that the method of education, not only with regard to the major fields
of knowledge, but also with respect to so-called character training, needs
to be critically examined and possibly considerably modified. Surely, what
is of utmost importance is that the mind of the young religious penetrate
fully the meaning of the truths presented to her; and this requires much
more than just the memorization of verbalizations and formulae. She must
be alerted to her habitual ways of evaluating, and helped to see discrep-
ancies between the latter and the principles she professes.
Since TAT Sequential Analysis can indicate the basic personality factors and point out discrepancies between theory and practice, this diagnostic procedure is undoubtedly a helpful technique with beginning classes of teachers. It would furnish invaluable information not only to those who are charged with instruction and counseling, but also to the young religious herself, with whom the task of developing the correct basic disposition lies.
CHAPTER VI

SUMMARY

The object of this investigation was to discover the characteristics that distinguish teachers rated by their pupils at the extremes in teaching efficiency. An examination of the methods by which previous research had approached questions of teacher efficiency suggested that the distinguishing characteristics of good and poor teachers could not be found in their intelligence and preparation alone, their external conduct as ascertained by behavior sampling procedures, nor in their responses to the ordinary personality inventory. It was felt that the differences lay below the external manifestations, and that their discovery required a disguised technique that was at the same time penetrating and objective.

The study was designed to locate differences between the teachers by means of the Otis Intelligence Test, the Minnesota Teacher Attitude Inventory, and the Thematic Apperception Test. The criterion groups were set up on the basis of pupils' ratings of their teachers on the Diagnostic Teacher-Rating Scale by S. M. Amatora. The subjects were three hundred teaching Sisters rated by their classes—a total of 10,720 pupils.

Intelligence test scores were found to be significantly different for the high- and the low-rated teachers on the secondary level, though not on the elementary. On both levels the younger teachers made significantly higher intelligence test scores.

The attempt to find the differences between the high and the low
teachers in the MTAI was partially successful in the grade school and not at all in the high school. The differences as revealed by this instrument lay chiefly in the teachers' attitudes toward the pupils and toward teaching. There were very few expressed attitudes, even in this area, on which the high-rated teachers consistently agreed.

The hypothesis that the differences could be found in the underlying personality structure by means of a projective technique was fully substantiated. By means of TAT Sequential Analysis an habitual basic disposition was found to characterize the high-rated group and to be plainly absent from the other. This basic disposition was seen as the habit of viewing life realistically and of responding constructively to its demands. The basic disposition was defined in detail in this study as it was expressed toward five common aspects of life: success, failure, loss, life and its obligations, and other people. The presence or absence of this disposition in the high and the low group respectively was verified by a predictive procedure of assigning the teachers to a group on the basis of their positive or negative scores by three judges independently.

Statistical procedures revealed the following specific results:

1. Intercorrelations of the seven areas on which the teachers were rated were comparatively high, ranging from .36 to .73 in the grade school, and from .14 to .69 in the high school.

2. Two questions of the rating scale, namely, Question II, on ability to explain, and Question VI, on amount of work assigned, correlated with intelligence scores of the high school teachers to a degree that was significant beyond the 1% level.
(3) There was no significant relationship between intelligence scores and rating in the elementary group.

(4) Mean MTAI scores showed differences in favor of the high-rated teachers that were significant at or beyond the 5% level only in the 7-8 grade group.

(5) When the items of the MTAI were categorized on the basis of their content, the 9-10 grade teachers were found to differ, at the 5% level of significance, in their opinions regarding pupil effort and discipline; while the 7-8 grade group differed similarly on the latter, and somewhat more significantly, at the 2% and the 1% levels respectively, on their opinions regarding pupil conduct and their liking for children and teaching. On both grade levels the differences were in the direction of higher MTAI scores on the part of the teachers who were rated higher by their pupils.

(6) Younger teachers, both elementary and secondary, made significantly higher intelligence scores than the older teachers (those above 41 in the grade school and above 47 in the high school), chi square being far beyond what is required for the 1% level of confidence.

(7) Younger teachers in the elementary school scored higher than older teachers on the MTAI, chi square again being significant beyond the 1% level.

(8) Chi square tests showed that age of teachers was no factor in the ratings of grade school pupils, but that it was an important factor, significant beyond the 1% level, in the high school, where the younger teachers were rated higher.

(9) When the teachers were paired so that the groups were similar in age and intelligence scores, and at opposite extremes in rating, the high
school teachers showed no differences on the MTAI, while the grade school
group differed significantly, beyond the 1% level, on their attitude toward
pupils and teaching, and somewhat less, at the 5% and the 2% levels respec-
tively, on their opinions on pupil conduct and on what should be done about
discipline. All of these differences were in the direction of a more favor-
able attitude toward pupils on the part of the high-rated teachers.

(10) Elementary teachers also differed significantly, at or beyond the 5% confidence level, on eight items of the MTAI, all but one of which had to do with their liking for pupils and teaching.

(11) When the TAT's of the teachers were scored according to a standard that had been empirically set up with forty cases, it was possible to dis-
criminate perfectly between the high- and the low-rated groups, both ele-
mentary and secondary.

(12) Scorer reliability on the TAT was 94.3% between two independent judges, and 97.2% and 96.6% respectively between each of the judges and the inves-
tigator.

(13) Differences between the TAT scores of the high-rated and the low-rated
groups of teachers were found by the Sign Test to be significant far beyond
the .1% level of confidence.

(14) All of the questions of the rating scale correlated significantly with the TAT scores, the highest relationship ($\rho = .96$) being between the scores and the pupils' liking for their teacher, on both elementary and sec-
ondary levels.

Briefly, the values of the present investigation can be said to consist
chiefly in the following: first, the definition of the necessary habitual
basic disposition of a religious teacher in terms applicable to the most significant aspects of her life; secondly, the demonstration of the effectiveness of TAT Sequential Analysis in revealing the presence or absence of that disposition; and finally, the assurance, within the limitations outlined above, that the pupils themselves can indicate by honest appraisal whether or not their teacher is motivated by the objective frame of reference here defined.
BIBLIOGRAPHY

I. BOOKS


Tyler, Fred T. "The Prediction of Student-Teaching Success from Personality Inventories." Berkeley, Cal., 1954.

II. ARTICLES


Amatora, Sister M. "Can Elementary School Children Discriminate Certain Traits in Their Teachers?" Child Development, XXIII (March 1952), 75-80.


Boardman, C. W. "An Analysis of Pupil Ratings of High School Teachers," Educational Administration and Supervision, XVI (September 1930), 440-446.


Bryan, Roy C. "Reliability, Validity, and Needfulness of Written Student Reactions to Teachers," Educational Administration and Supervision, XXVII (December 1941), 655-665.


-----. "Efficiency of the MTAI for Predicting Interpersonal Relations in the Classroom," Journal of Applied Psychology, XXXVII (April 1953), 82-85.


Dodge, A. F. "What are the Personality Traits of the Successful Teacher?" Journal of Applied Psychology, XXVII (August 1943), 325-337.


Drawhorne, Curtis L. "Relationship Between Pupil and Student-Teacher Interaction and Pupil Ratings of Teaching Effectiveness," Educational Administration and Supervision, XL (May 1954), 283-296.


------. "Specific Traits and Teaching Competence," Educational Administration and Supervision, XXXIX (December 1953), 500-503.


------. "A Second Attempt to Predict Teaching Success from Word Fluency Data," Journal of Educational Research, XLIX (September 1955), 13-25.


Laycock, S. R. "The Bernreuter Personality Inventory in the Selection of Teachers," Educational Administration and Supervision, XX (January 1934), 59-63.


-----. "Superior and Good Teachers," *School and Society, LXXXV* (October 26, 1957), 315.


APPENDIX

AN ILLUSTRATION OF TAT SEQUENTIAL ANALYSIS

The following is the Sequential Analysis of the TAT stories of a high-rated teacher, aged 23, with a Gamma IQ of 126. The statements are scored according to the Scoring Standard set up in this study. The stories are presented in the succeeding pages.

Sequential Analysis

1. When they make you do some work that you don't like to do, you can find a way to escape from it to the pleasure you want. *C -

2. But when you see the sacrifices of others, you admire their strength and realize that that is to be found somewhere other than in books. D +

3. Still, you eagerly await what you are looking for; and when you are disappointed, you feel bad. C -

4. But if you wait long enough in spite of the suspense, the good news will come. C +

5. When people overcome their fears and look into the cause of them, they find out that there is no need to fear. A +

6. It may take you a while to overcome your selfishness when you are asked to make a sacrifice graciously, but after thinking about it, you do it. C +

7. It takes strength to relinquish what you want and to give in to what you see is right. C +

8. Sometimes there are better things to do than those that you first thought best. C +

9. You may think your plans to escape your obligations are clever and working well, but they may be frustrated in a way you do not expect. D +

10. And a loss may bring about a good result. C +

*Letters refer to categories.
11. What you wanted may have been given to someone else, and as you try to rectify your mistakes, you have the consolation at least of knowing that you can no longer be deprived of it as punishment for them.

12. You may fear that your surroundings will warp you, but if you are there long enough you will be made into a creature capable of winging your way to the place you want to be.

Positive Score = 10.

The Stories

1. Timmy was an only child, and his doting mother wanted to fashion her darling according to her own ideas of what he should be. One of the things Mother most wanted him to do—and he wanted least—was to take violin lessons. Timmy's friends had organized a baseball team, and were waiting on the sidewalk near his home for Tim to come out one Saturday. Tim heard the whistle as he was sitting glumly looking at the hated fiddle, and his mother's injunction that he must practice two hours this afternoon ringing in his ears. "Don't see any use in practicing on this old thing!" Tim was grumbling to himself. "No wonder Jack's always calling me a sissy the way I hafta do such girls' stuff!" Then his face brightened as he spied the open window. A few seconds later, the gang and Tim were racing toward the ball park.

2. The crops had been poor that summer—no rain for almost six weeks. As May stood near the field, school books in one hand, she wondered whether her Mom knew the thoughts that had been whirling in her head for days. Now, on the opening day of school, she had to let them out. "Mom, are you sure you and Joe can afford to let me go? I mean, Dad's death and all, and now this summer—" "May, if it's the last thing I do, you're going to get a chance to finish high school." As May walked toward the road, she glanced back at her Mother, tired looking, and yet—"I hope someday I have the same sort of strength Mom has, and I know I won't get that from books."

3. She'd been waiting all week for him to call. The prom was a few days away, and she'd been so sure he'd ask her. When it finally rang, she ran the steps two at a time, settled herself on the floor with the phone, and prepared for a long conversation. "This is Harrington Dry Cleaner's" came a voice. "Will it be all right if we deliver your things this afternoon?" "Yes," she choked, and hung up. Then, though she knew she'd be just as eager the next time the phone rang, she buried her head on her arm and sobbed.

4. They'd lost one baby at birth. Joe, their second, would also be their last the doctor had said. As Joe, Sr., and Alice waited in the hospital
corridor, Alice said, "The nurse should be out soon now." "How long does it take them to pump that stuff out of his stomach anyway?" "Well, he drank almost the whole bottle of furniture polish, dear." "I can't stand the suspense any longer. I'm going in." Alice clung to him, "Please, Joe, that won't help any of us." Then the door opened, and a smiling nurse announced, "You can come in now. Your little boy will be just fine."

5. Aunt Amy was the dearest person I ever knew. She lived all alone in a big rambling house now that her three children were all married. I used to ask her if she was afraid ever at night in that large house. She'd only laugh, and ask what I thought she had to be frightened of. Then one night when I was staying with her, I heard some strange noises shortly after midnight. I tiptoed out to the hall, and met her coming out of her room. She looked--well, not frightened maybe, but a little different than usual. I followed her downstairs, not daring to even whisper. As we stood in the lower hallway, the sound came again--the piano in the parlor. Aunt Amy started toward the door. "Don't," I said. "Call the police." But she swung open the door and peered in. Then she switched on the light. As she broke into laughter, I rushed in to see Dinah, the black cat, standing on the ivories and blinking at us.

6. "Mother," John pleaded. "Please come to the station with me." His mother continued to stare out of the window in stony silence. "All right," he sighed, "have it your way. But you're awfully selfish!" The door slammed after him. She stood there long minutes, thinking of John, her last and favorite son, and of Mary. Mary--the girl who would take him from her. Mary, who was coming now on a week-end visit. She knew John had hoped she would like Mary, but he didn't understand how lonely she'd be when he was gone. Selfish! She thought of his baby days, high school, the hopes and plans she had for him--and of his father, whom he resembled so closely. At last, she gave a tired sigh and turned toward the kitchen. Then she quickened her step as she planned the lunch she'd have ready to welcome Mary home!

7. She looked across at them--the white head of the old man so near to the darkness of his son's. "And just as different as the color of their hair are their ideas," she thought. "And yet they're so much alike. The same stubborn jaw!" They had been silent for several minutes, both looking into the blazing fire. Neither seemed to remember she was present. Then her son spoke. "Dad, if you really would--" "My boy," the older man interrupted, "if you really want to go, then all I can say is--be the best priest you know how." Then, as she saw them both smile a little, she thought, "No, not the same stubborn jaw, but the same quiet strength."

8. "Draft dodger," they called him in September when he went off to college, and their sons to a training camp. But Bob didn't mind. Probably because he didn't know. Then as the years went by, and the war continued, he couldn't help hearing. He tried not to become bitter. "They just don't understand," he'd tell himself. Then finally it was graduation day, and
as he held his diploma in his hand he was dreaming his favorite dream--
Dr. Bob, army surgeon, serving his wounded countrymen. Months of intern-
ship, enlistment, and then Iwo Jima. Then one day came a letter that made
it all seem really worthwhile. "I want to apologize," it read, "for hav-
ing referred to you as a draft dodger. You have saved my boy's life, and
taught me there are other ways to fight wars than with machine guns. Thank
God for your surgeon's knife!"

9. It was hot and dusty, and they'd been walking since late last night.
It hadn't been easy sneaking out of the Boys' Reformatory, but thanks to
Jack's clever planning, they'd managed. Now here they were, across the
county line already, and it was only nine o'clock. "Let's rest awhile," said
Jack. "There's a creek over there." They slipped under the fence, across
the field, and were soon lying on their backs near the cool water. It was
only a matter of minutes before they were asleep. The sun was overhead
when Bill woke with a start. There, staring right at him was the biggest,
maddest bull he'd ever seen. "Hey, fellas," he hollered, and they started
across the field with the bull in hot pursuit. "The trees!" yelled Jack.
"Climb one of the trees!" And so they did. It was there the police found
them, hungry, frightened, and angry at having their master-mind plans frus-
trated by a bull.

10. It had been years since they'd met--these two brothers. Years of sep-
aration and disagreement. Now their mother had died and they were coming
for the funeral. Tom arrived first, and was standing near his mother's
coffin, when Dan walked in. The others in the funeral parlor watched with
bated breath as Dan walked over next to Tom. Tom turned--they looked long
at each other--then quietly and swiftly embraced each other. Many a mourn-
er shed tears of joy as they watched the brothers kneel at their mother's
coffin together.

11. A chocolate cake stood on the table, warm, inviting, adding a final
touch of hominess to the kitchen. But Jan didn't feel very homey in that
kitchen right now. As she looked at the cake, her mouth watered. But as
she looked at the puddles she had dripped from the back door to the table,
her mouth got very dry in anticipation of what would happen. "Hello, dear,"
said Mother as she came into the kitchen. "Hello," croaked Jan, waiting
for Mother to see. But Mother apparently didn't notice. Then she scooped
up the cake, smiled, and said, "This is for Mrs. Barnabee. I'll take it
over to her now." As her mother went back into the living room, Jan did the
best she could at wiping up the muddy tracks, and thought, "Well, at least
my punishment won't be no cake for dessert."

12. It was very dark in the ruins of the bombed section of Milan. Crawl-
ing along on a wide ledge, a caterpillar was the only sign of life in the
desolate scene. "A caterpillar doesn't mind going through life alone," he
mused, "but in such squalid surroundings, what a warped attitude he can get.
How I'd love to be out in the green fields that mouse was telling me about
yesterday." At the edge of the wall, there was an opening where the sun came through. Here the little caterpillar sat, and thought his long, philosophical thoughts all day—and at evening an old man passing along the street stopped to admire a dainty butterfly gracefully winging its way over the ruins, toward the fields at the edge of the city.
APPROVAL SHEET

The dissertation submitted by Sister Mary Innocentia Burkard, SSND has been read and approved by five members of the Department of Education.

The final copies have been examined by the director of the dissertation and the signature which appears below verifies the fact that any necessary changes have been incorporated, and that the dissertation is now given final approval with reference to content, form, and mechanical accuracy.

The dissertation is therefore accepted in partial fulfillment of the requirements for the degree of Doctor of Philosophy.

May 2, 1918

Date

Signature of Adviser