1932

A study of teachers trained at the College of the Pacific in regard to mental ability, scholarship, and teaching success

Emma Pearson Fentzling
University of the Pacific

A STUDY OF TEACHERS TRAINED AT THE COLLEGE OF THE PACIFIC IN REGARD TO MENTAL ABILITY, SCHOLARSHIP, AND TEACHING SUCCESS

by

Mrs. Emma Pearson Fentzling

May, 1932

A Thesis

Submitted to the Department of Education

College of the Pacific

In partial fulfillment

of the

Requirements for the

Degree of Master of Arts

APPROVED:

Head of the Department

DEPOSITED IN THE COLLEGE LIBRARY:

Librarian

DATED:

To

My Parents

CONTENTS

LIST OF TABLES

Number Page

LISTS OF CHARTS

CHAPTER I

INTRODUCTION TO THE PROBLEM

It is only within the past two decades that scientific measurements have been applied to the evaluation of teaching success. Elliott in 1910 composed a rather detailed rating scale of 100 points. This scale was the basis of much experimentation but was found to be too complicated for practical use. Several years later appeared Boyce's score card with ratings on 45 points. This was more adapted to evaluating teaching success. Another interesting method of rating personality was invented by Walter Dill Scott. His method employed comparisons of several individuals on what was termed the man-to-man, five-point scale. This scale was used to rate officers in the army. During the past decade many other rating scales, the first important factor in the evaluation of teaching success, have been devised. The status of these scales in the field of education will be discussed in Chapter III.

A second important factor in the evaluation of teaching success other than by the rating of personality traits is mental testing. Intelligence testing is the resultant of at least five converging movements: namely, Binet's experimentation with the feebleminded, Wundt's and Ebbinghaus' laboratory tests in experimental psychology, Cattell's and Thorndike's study of individual differences, the Galton-Pearson development of statistical procedure, and the Galton's, Wood's, Cattell's studies in anthropology.[1] However, the use of

[1]See P. M. Symonds, Measurement in Secondary Education, 53 - 55.

mental tests in the schoolroom have become prevalent owing to the successful experiments with such tests during the World War.

The third factor commonly measured in either predicting or evaluating teaching success is scholarship. This factor would probably prove more reliable if it were based on standardized tests such as those made by Ruch, Toops, Pressey, Seashore, and others. The factors of intelligence tests and college grades are considered in Chapter II.

PROBLEM

The successful teacher possesses certain indispensable personal traits and professional attitudes. Are these traits and attitudes associated with high intelligence? Are high intelligence test scores predictive of success in teaching? Do successful teachers have a high degree of scholarship? What is the relation of a high college scholarship to success in teaching? Can the major professors predict success in teaching on the basis of success in college? How accurate is the critic teacher's report in predicting professional ability? Finally, what personal traits and professional attitudes are considered the most important by the principals and how are teachers rated on these? Consideration will be given to these questions to find out what constitutes success in teaching as indicated in the principals' and superintendents' reports.

LIMITATION OF THE PROBLEM

The classes selected for this study are those for which complete measurements are available in the offices of the College of the Pacific.

Four specific factors are used in the investigation: namely, (1) intelligence test scores, (2) college grade-point-averages, (3) ratings predicting teaching success given by the major professors and the critic teachers, and (4) principals' ratings on teaching success. The students in the entering freshmen classes of 1923, 1924, 1925, 1926, and 1927 who upon graduation from the College of the Pacific entered the teaching profession are selected for particular consideration. Of the 867 freshmen of this five-year period, about 175 have taught or are now teaching. However, at the time of this study records are complete for only 122 teachers.

METHOD OF SECURING DATA

The data used for this study are intelligence test scores, college grade-point-averages, major professors' ratings on personality, critic teachers' ratings on professional ability, and principals' composite ratings on both personality and professional traits.

The test scores were secured from the Thorndike Examinations or the Thurstone Psychological Tests given to all entering freshmen of the College of the Pacific in the classes from 1923 to 1927 inclusive.

The grade-point-average for college scholarship represents the total number of honor points divided by the total number of units for which the student is registered. An A equals 3 honor

points, B 2 honor points, C 1 honor point, D none, and E -1 honor point.

Of the several teachers recommending the student for a teaching position, the major professor's rating is selected because he, presumably, has been in closer contact with the student for several years and has had more time and opportunity to observe the student's particular personality traits. Although given the general heading of personality, this rating includes the following items: personality, scholarship, judgment, health and vigor, personal appearance, power of oral expression, energy and persistence, culture and refinement, and community interest and standing. (See appendix for form of the rating scale.)

In many instances only one critic teacher reported on the work of the prospective teacher. Therefore, only one critic report is selected for this study, and the items are selected which rate professional ability: namely, skill as an instructor, ability in discipline, and influence on the students. This rating is based on one semester of the student teacher's work.

Since the period of teaching experience for this group may range from four and a half years to only one half year for the last graduating class, it was decided to use only one principal's rating report for each teacher. Where more than one report is available the last one is used. A composite rating is made from each principal's report, which includes all the previous mentioned items of personality and items of professional ability. Then the specific items of skill and discipline are compared to the composite rating.

TREATMENT OF THE DATA

To determine the value of intelligence test scores, scholarship, and predicted success with teaching success, the survey method is employed.

All measurements and ratings of the 122 selected teachers are divided on a five-group basis. Only one intelligence test score is used; only one scholarship record is used, that which is based on the four-year college course; and 3 separate ratings on teachers' traits are used.

The five groups for each measurement are not divided on an equal basis. The type of measurement and the distribution of numbers for that measurement suggest the basis for the five divisions of each. Group 1 in intelligence test score is compared with Group 1 in scholarship and both are compared with teaching success. Many other similar comparisons are made.

Because one measurement is objective and another subjective, and because one measurement is on a true quintile basis whereas another is on an arbitrary basis, the findings of this study are not presented as showing exact comparisons between the various factors in a scientific relationship. The purpose of the study is to bring together those available measurements which present-day research has found more or less reliable in determining success in teaching. Only in a general way are the various groups related to one another and the degree of

teaching success is determined.

SUMMARY

This study of teachers trained at the College of the Pacific in regard to mental ability, scholarship, and teaching success deals with the analysis of intelligence test scores, college grade-point-averages, major professors' and critic teachers' recommendations predicting success, and superintendents' and principals' ratings on teaching success. Intelligence test scores are an objective measurement of mental ability. Unless teachers' marks are given on the basis of standardized tests, grade-point-averages are subjective measurements based on personal opinion. The popular rating scales in current use are also subjective measures. However, this present study is justified because of the fact that "research into teacher rating and personnel management indicates that the best criterion now in use for determining the relative effectiveness of teachers is personal opinion".[1]

[1] Evelyn Clement, "An Evaluation of Teacher-Training", Educational Administration and Supervision, Feb. 1932, p 91.

CHAPTER II

SOME RECENT INVESTIGATIONS ON THE PREDICTION OF TEACHING SUCCESS

As a background for this study of teachers trained at the College of the Pacific, several opinions will be given on the results of the prediction of teaching success.

To enumerate even in brief form the many different factors which have been segregated in the past few years in an attempt to determine their relation to future success would be an impractical task. Leading educators have conducted researches with individuals and classroom students, in many grades and classes, and even with certain groups throughout the entire educational system. Often records are available for the student's complete school performance and for his post school performance. Graduate students as well as administrators everywhere seem interested in this problem of predicting success, if one were to judge by its frequent mention in all types of educational literature. This study is concerned only with the prediction of success on the college level particularly as it pertains to future teaching success. The methods used will be illustrated with citations.

Among the several criteria for predicting teaching success in use at the present time, three seem to be most commonly employed; namely, intelligence test scores, scholarship, and personality rating scales. The data of this chapter deal with the prediction of success in relation to teaching success as shown by tests, grades, and recommendations. One of these criteria may present more reliable

conclusions than another; yet no one measurement is as reliable as a combination of two or more. For the study of the prediction of teaching success, all known and available student information should be considered.

INTELLIGENCE TEST SCORES

Wood summarizes some important uses of intelligence test scores as follows:

> The intelligence test is the peerless admission criterion when used in conjunction with other available criteria. In the matter of predicting academic achievement, the intelligence test stands without a serious rival. By the use of the intelligence test many candidates who have made satisfactory or passing grades in high school, but who have not the superior intellect requisite to success in college, are saved from discouraging failure and waste of time and money by being advised to undertake work more suited to their talents...The use of the intelligence test makes it possible to select the very superior minds among entering students for early attention... One of the greatest advantages of the intelligence test is that it enables the administration to adjust academic load to individual capacity so much more precisely than heretofore that many mediocre students, who would otherwise have dropped out of college, are saved from failure during the first year, and thus are allowed to go on through college with as much speed and profit as their intelligence allows.[1]

Intelligence tests do not pretend to measure an individual's particular talent but rather to measure mental capacity in a general way. Burr in the following citation suggests the limitations of tests.

> Experiment after experiment has shown that the I. Q. correlates most closely with accomplishment in the following subjects: vocabulary, reading, spelling, and arithmetic in the elementary grades, and Latin, English composition, and higher mathematics in the secondary schools. Many tabulations have shown a noticeable lack of correlation between the I. Q. and accomplishment in music, the fine arts, the industrial arts, and physical education.[2]

[1] Ben D. Wood, Measurements in Higher Education, 274.

[2] S. E. Burr, "Why the I. Q. Needs a New and More Descriptive Name," The Nation's Schools, Apr. 1321, p 52.

Breckenridge reports a study on the predictive value of high school averages and test scores for Normal school freshmen. The principal of the Louisville Normal School made a study of all its graduates from June, 1921, to February, 1927, inclusive. The results of the various lines of investigation lead to the following conclusions:

1. A student's record in high school is, to a significant degree, prophetic of her subsequent record in the subject-matter studies in normal school.

2. The high school record is of considerably less value in predicting success in practice teaching, particularly where the practice teaching is rated by principals in a city system.

3. Intelligence, as measured by the Army Alpha Test, is to a slight degree predictive of achievement in normal school subject-matter studies.

4. Intelligence tests are of little or no value in predicting success in practice teaching.

5. The combined action of high school marks and intelligence test scores has greater predictive value than either criterion taken separately.

6. The closer relationship found to exist between high school grades and normal school subject-matter marks, student-teaching grades and cadet teaching grades shows that the elements of perseverance, initiative, interest, attitude, etc. included in all grades are more closely related to teaching than are intelligence scores.[1]

Another writer expresses more confidence in intelligence test scores in predicting the success of the prospective teacher. Gist, of San Francisco State Teachers College, suggests that no individual should elect teaching whose I.Q. is much below 90. He also suggests that an individual with a very high I.Q. should be placed with great care in a teaching position, because "the genius or the near genius might be so unsympathetic with the slowly developing mind as to be entirely unsuccessful."[2]

[1]Elizabeth Breckinridge, "A Study of the Relation of Preparatory School Records and Intelligence Test Scores to Teaching Success," Educa. Admin. and Superv., Dec. 1931, 641 - 648

[2]Arthur S. Gist, "Important Points of View in Teacher-Training", Educa. Admin. and Superv., April 1931, 269.

Freeman does not agree with Gist in setting an arbitrary standard below which prospective teachers should not be accepted. He believes that one or more of several different factors may influence the intelligence test score. He interviewed 68 sophomores in regard to intelligence test scores and college grades. Forty-two students who ranked in the upper 30 per cent on the intelligence tests, averaged less than 75 in their courses of the freshman year. The reasons which they gave for failures in study were as follows: lack of interest, poor habits of study, loafing, athletic competitions, extra-curricular competitions, work for self-support, reading and study outside of courses, social activities, no apparent reason, and illness. Twenty-six students who ranked in the lower half on the intelligence test made freshman scholastic averages which placed them in the upper 30 per cent of their class. The reasons which they gave for their poorer performance on the intelligence test were as follows: newness of conditions and nervousness during first days, attached no importance to test, handicapped by time limit, illness, and no apparent reason. Freeman reaches this conclusion:

> The correlation technique. . . is of unquestionable assistance in studying predictive values; but it should also be clear that there are sufficiently frequent elusive factors which make it necessary, for a better understanding of a tests' predictive value, to supplement the objective data with subjective reports in those instances where discrepancies exist between test rank and course grades.

A similar study was conducted by Hughes for 57 students of the 1923 graduating class of Pasadena High Schools who entered three different universities. Only one table is reproduced here to show the relation

[1]Frank S. Freeman, "Elusive Factors Tending to Reduce Correlations between Intelligence Test Ranks and College Grades", Sch. and Soc., XXIX, No. 755, pp 784 - 786.

between the mental test and the first semester college marks. Student Number 1, highest in intelligence score was also considered by teachers and students to be strong in the personality traits that make for success in study. Student Number 9 also possessed those traits.

THORNDIKE INTELLIGENCE SCORES IN RELATION TO UNIVERSITY ACHIEVEMENT

(9 Stanford Freshmen)

Thorndike Scores	Students	Correlation 0.24	Students	Achievement Ratios
93.2	1		1	2.00
91.8	2		3	1.65
89.5	3		9	1.56
83.4	4		2	1.18
82.5	5		5	1.10
68.1	6		7	1.06
61.5	7		4	.97
58.7	8		6	.96
57.2	9		8	.75

Hughes' comments on the table are as follows:

Now, it would be surprising to find a close agreement between intelligence scores and school success when some students of average ability are using that ability to the maximum degree while other students of extraordinary ability are loafing on the job and acquiring habits and attitudes incompatable with success. Add to this fact that school marks are extremely deceptive, sometimes standing for actual academic achievement but more frequently, perhaps, representing a combination of more or less desireable personal characteristics, and we have a sufficient explanation for such discrepancies as we have charted above. But even if instructors' marks were dependable measures of

school achievement, and intelligence scores were absolute gauges of human ability, we should still find, under present conditions of school, home, and community, many reasons for lack of agreement between ability and academic success.[1]

Another specific example of the predictive value of intelligence quotients and their correlation with other factors is found in the work of Gillis in her study of four recent graduating classes of Barnard College. She found the following:

(6) No large differences in median intelligence scores are found among the four groups of subjects in which students may major--humanities, social science, physical science, and biological sciences.

(7) Certain subjects, as physics, anthropology, mathematics, zoology, psychology, and English seem to select students of superior ability.

(8) Students majoring in art, music, and the classics make median scores which are appreciably lower than the median for their group. This may be explainable by the fact that interest and excellence in these subjects may depend more upon the emotional make-up than upon the intellectual quality of the individual.[2]

Teachers are prone either to pride themselves on the popularity of their courses which attract many students or to pity themselves because of the difficulty of the subject matter in their rather obsolete courses. It is recognized that certain major courses require a high degree of intelligence, but that does not mean that an individual of lower intelligence has no chance to succeed in that subject. Investigations have shown that each major subject generally includes individuals whose test scores range from the highest to the lowest score.

[1]W. Hardin Hughes, "Why Intelligence Scores are not more Highly Predictive of School Success", Educa. Admin. and Superv. V XII, No. 1, pp 44 - 48.

[2]Frances M. Gillis, "Correlates of Intelligence in College Students", School and Society, Aug. 22, 1931, p 270.

Dexter of Agnes Scott College conducted a survey of 4 representative colleges on the subject of intelligence test score and major subject. She concludes:

> About the only fact standing out conspicuously from this table is that students selecting English rank highest on the whole.
>
> It is concluded that there is a considerable degree of variation among colleges as to subjects attracting the brightest students, with no conspicuous advantage in favor of any one subject. Each institution seems to offer greater attractions in some departments than others -- perhaps a matter of temporary popularity. Furthermore, given intelligence and training enough to get into college, a student can succeed as well in one subject as another, provided he cares to do so.[1]

COLLEGE MARKS

As long as schools have existed there has been a measuring device of one kind or another applied to the students' work. Marks or grades as commonly used should mean only one thing and that is achievement. Other devices may be employed to give recognition to natural ability, studiousness, speed, neatness, accuracy, personality and the like. Most teachers are influenced by these other traits and give grades on the general impression of the student rather than on his accomplishment, thereby giving a higher or lower grade to the individual than he actually deserves. One college dean found that over a period of 11 semesters 38 to 49 per cent of all the marks given by the faculty were A's and B's. He concluded that either there were more students of genuine attainments than they were willing to recognize when they talked shop with each other, or they were rewarding with scholar's marks people who were in nowise scholarly.[2]

[1]Emily S. Dexter, "Intelligence - Test Score and Major Subject", School and Society, V. XXX, No. 780, pp 779, 780.

[2]A College Dean, "Concerning Marks", School and Society, Oct.10,1931, P. 506

Middlebrook reports a study for about 2000 students whom she taught during a period of 7 years. Her records of grades represent the distribution of the normal probability curve. However, from some follow-up cases she concludes the percentage passed was too high. She suggests that no students with grades below 70, or a "C", be allowed to enter college where they will be influenced by professional ideals and aspirations wholly incompatible with their mental capacities. They thus create hopes for themselves the fulfillment of which it is as impossible as it is undesirable.[1]

Of late years, these two factors, intelligence test scores and school marks, have had considerable influence in the selection and guidance of students. However, the many experiments with thousands of students have made apparent the need for the measurement of a third factor, personality.

PERSONAL RATING SCALES

In a study of 357 selected students entering the State Teachers College, at St. Cloud, Minnesota, in 1925, McCrory made an attempt to determine which students applying for admission to the college are likely to make an unsatisfactory academic record (less than a "C" average). One conclusion is as follows:

> For prediction of success in college no one criterion alone is reliable. Intelligence test scores and high school records offer at present the most dependable data. It is quite probable that we have gone as far as we can go in psychology in predicting future success on the basis of things intellectual. The next development in this

[1] L. Ruth Middlebrook, "A Modest Proposal", School and Society, Oct. 10, 1931, p 507.

field must come through experimental study of character and personality traits. There is need for experimental determination of the exact relation to success of specific character traits such as interest, initiative, application, cooperation, sincerety, etc.[1]

Test scores, grades, and personality traits are compared in different combinations in the following study of Neel and Mead.

A select group of 64 college seniors of Ohio Wesleyan University preparing to teach in high school furnish significant data for correlations between certain group factors. The correlations as found are as follows:

1. General scholastic average with student-teaching	⁄.369
2. Percentile rank (mental ability) with student-teaching	⁄.141
3. Personality traits (Almy-Sorenson) with student-teaching	⁄.689
4. Achievement in subject-matter with student-teaching	⁄.486
5. General scholastic average with subject-matter	⁄.753
6. Personality traits with general scholarship	⁄.291
7. Mental ability with general scholarship	⁄.486

. . .The study was undertaken to ascertain the simple correlation between achievement in supervised student-teaching and the other group factors. In view of the findings, the effective factors seem to be the following, in the order given:

1. Rank in selected group of personality traits;
2. Rank in subject-matter achievement;
3. Rank in general scholastic achievement;
4. Rank in general mental ability.

It is significant that the highest correlation with status in student-teaching is that of status in a selected group of personality traits, not complete personality.[2]

Bowman of De Pauw University has suggested three principal methods of estimating probable teaching success other than those of intelligence test scores and college marks. Following is the summary of his points.

First there is the use of interviews, letters of recommendation, and rating scales. The second is that of judging the teacher in the light of observed activities on the part of his pupils. The third is

[1]John R. McCrory, "A Study of the Relation between Ability and Achievement", Ed. Admin. and Superv., V. XII, No. 7, pp 481, 490.

[2]Mary O.Neel and A. R. Mead, "Correlations between Certain Group Factors in Preparation of Secondary School Teachers", Educa. Admin. and Superv., Dec. 1931, pp 675, 676.

the testing method. This method, if it can be made successful, would possess the double advantage of being more objective and more easily administered than the other two. It is very difficult to secure eight independent ratings on many people and it is even more difficult to place the prospective teacher in an actual teaching situation early enough and for a long enough time to get an estimate of probable success which would be of much help in solving the problem of selection of desirable candidates for our teacher-training courses.

. . .There are six important phases of the testing approach to this problem: namely; (1) knowledge of the child to be taught; (2) knowledge of the society of which the child is a member; (3) specialized training in some field of knowledge; (4) knowledge of the school system; (5) teaching technique; (6) personality. We have some data with respect to the significance of each of these six general items in teaching success. The coefficients of correlation between teaching success as estimated by superintendents' and principals' judgments and grades or scores made in academic subjects, in general courses in education, in student teaching, and in intelligence tests have not been high. In fact, many of them have been quite low. Whitney found the multiple correlation of all these factors plus physique, with teaching success to be but .288. In no case does Tiegs report a coefficient of correlation between the selective device and the criterion higher than .27. The low coefficients of correlation may be explained by the unreliability of the criteria [subjective judgments of principals, tests of scholarship and intelligence.]

He concludes that since we have neither a satisfactory measure of predicting such success, even as we now define it, that fact

. . .points to the conclusion that we must use every scrap of available evidence in making our selections for teacher-training courses, but it does not indicate that we should do nothing about the matter. As little above chance as our means of prediction are as applied to the early selection of teacher-training candidates, they are probably about as good as any that can be applied after the teacher-training is over, at least with the possible exception of student-teaching grades.[1]

One further example of personal rating scales which specifies their content and the results of their use in predicting teaching

[1] Earl C. Bowman, "The Problem of the Early Prognosis of Teaching Success", Educa. Admin. and Superv., Feb. 1931, pp 95 - 102.

success will suffice to summarize recent investigations and studies conducted on this phase of educational progress. Armentrout presents the results of a comparative study of the ratings of 200 teachers by training teachers and public school superintendents. The training teachers rated their teachers during their period of student-teaching at the Colorado State Teachers College, and the superintendent rated them during the year following their graduation from this institution. The same rating card of 16 characteristics to be checked on a five-point scale was used. The results were as follows:

> The eleven traits rated higher by the training teachers are as follows: interest in life of the community, ability to awaken interest and effort, ability to get on with pupils, control, willingness to co-operate, desire for professional growth, loyalty, leadership, initiative, interest in life of school, and psychological method.
>
> The training teachers rated scholarship, correct use of English, instructional skill, voice, and originality lower than do the superintendents. The training teachers come into more intimate contact with these traits than do the superintendents and perhaps have a higher standard for judging all of them with the possible exception of voice.
>
> The evidence is quite clear that both training teachers and superintendents rate too high; there are too few C and D ratings and too many A and AA ratings. There are more than twice as many A's among the superintendents' ratings as in a normal distribution.[1]

SUMMARY

Recent investigations on the predictive value of intelligence tests have shown that tests predict academic achievement but do not predict success in student-teaching. Neither are grades predictive of success in student-teaching. Ratings on a selected group of

[1] W. D. Armentrout, "The Rating of Teachers by Training Teachers and Superintendents", The Elementary School Journal, V. XXVIII, No.7, 511 - 516.

personality traits predict a closer relation with success in student-teaching than does any other single factor. Intelligence test scores and grades have predictive value when combined with personality ratings. These ratings, however, are subjective and they show a decided distribution toward the high ratings.

CHAPTER III

SOME RECENT INVESTIGATIONS ON THE RELIABILITY OF RATING SCALES FOR THE EVALUATION OF TEACHING SUCCESS

The factors of intelligence test scores and college grades are not sufficiently reliable to predict with any degree of accuracy teaching success as shown by recent investigations discussed in Chapter II. Neither do the other factors of age, sex, nationality, social level, etc. predict teaching success. However, there is one factor which educators have devoted much attention and research to particularly since the above mentioned factors have not contributed to the solution of the problem. That factor is personality. Among the first experiments in this field were those by Professor June Downey.[1] Other scales for the testing and rating of character traits have appeared for children as well as for adults and for industrial and business as well as for professional levels. At present these scales enjoy the most popularity and probably the most reliability in evaluating success.

The data of this chapter deal with those certain personal traits that are recognized as being indispensable to successful teaching. In October, 1928, Peterson and Cook collected data from state teachers' colleges, state Normal schools, and city

[1]June E. Downey, The Will Temperament and Its Testing.

Normal schools, representing 45 states, on score cards and rating devices for evaluating student-teaching. Twenty to 30 or more items were usually listed on these sheets for the supervision and the final evaluation of student-teaching on a five-point scale of merit. The custom is about equally well established of placing full responsibility for rating on one individual, and of dividing it between two. The main items of the Composite Rating Scale for general efficiency are personality, general preparation, professional attitudes, teaching qualities, management, results of teaching, and extra-classroom activities.[1]

Much valuable information has been obtained by having students rate their teachers. In June, 1927, Newmark had 223 students analyze their elementary or secondary teachers for the best teacher and the poorest teacher. He found the characteristics frequently mentioned by these students in the Philadelphia Normal School were the same indispensable traits of the teacher mentioned by Bagley, Knight, Jones, Withers, Morrison, Anderson, Bird, Palmer, Book, Elliott, Burton, Ruediger, and Rugg. Among these characteristics are the ability to discipline pupils, pleasing personality, enthusiasm, scholarship, health, sympathy, good character, personal neatness, good voice, command of English, and sincerity.[2]

[1]See Oda K. Peterson and Wm. A. Cook, "Score Cards and Rating Sheets in Teacher Training", Educational Method, V. IX, No. 6 322-330.

[2]See David Newmark, "Students' Opinions of their Best and Poorest Teachers", The Elementary Sch. Jr., V. XXIX, No. 8, pp 576-585.

Rating scales, if they are to be effectively employed for comparative purposes, must be used by trained and competent individuals. Dr. Walter Dill Scott, an authority on this subject, suggests that each quality on the scale be defined, and the definition alone be used as a means of identifying the quality.[1]

THE RELIABILITY OF SINGLE VERSUS MASS OPINION

Mass opinion, or the pooling of at least 8 independent judgments,[2] is considered to be as reliable as it is possible to obtain. However, Conrad in a recent review of a study states that only one rating by a single individual compares favorably with the several ratings by several other individuals. He says that the army psychologists studied the ratings given by at least 5 different rating officers on the men in 5 companies at Camp Meade. These ratings were adjusted on the basis of the normal curve, and were then thrown into a single contingency (or correlation) table. This attempt to remove the personal equation made very little difference in the correlation between the ratings and test scores, and it led to no appreciable improvement in the ratings.[3]

There is general agreement on the varying judgments of different individuals; but it is not usually understood that the individual

[1] See W. H. Hughes, "Refining the Estimates of Personal Qualities," The Nations' Schools, Feb. 1931, p 55.

[2] See P. M. Symonds, Measurement in Secondary Education, 354.

[3] See H. S. Conrad, "The Effect of the Personal Equation", Jr. of Ed. Psychology, Feb. 1932, P 147.

varies in his ratings on the same scale for the same person. This would appear to be a strong argument in favor of many opinions, particularly when ratings are more or less indefinite. The following experiment illustrates findings opposite to those reported by Conrad on the reliability of individual ratings.

In 1918 Rugg investigated in a most thorough manner the reliability of the Army Rating Scale. He had opportunity to analyze the experimental situation made possible by the World War. Nowhere else had such carefully controlled factors presented themselves for such scientific study. For instance, there was a group of 461 very intelligent army officers whose average Alpha score was B+ who cooperated in the construction and the criticism of rating scales. These officers were in constant association with one another for about a year. They were therefore qualified to give accurate reports on observations of one another. Two or more official ratings which were made on each officer were compared on an 80 point scale for about 2,383 cases. The following table shows that one rater varies almost as much in his two separate ratings on a lieutenant as different raters vary in their separate judgments on the same individual rated by them.

	Same Rater	Different Raters
For second lieutenants	10.2 points	12.0 points
For first lieutenants	10.2 points	21.7 points
For captains	8.4 points	16.9 points

Rugg continues:

About half of the differences were increases and half decreases. The medians were somewhat different, but the general conclusion was inescapable: it was very improbable that an officer was located within even his proper "fifth" of the entire scale by an "official" rating. The ratings were perfectly useless. (And these rating conditions are quite comparable, if not superior, to those of education--certainly as to education and experience of raters, administrative control over rating and the like.)[1]

A similar experiment of rating by superior officers was conducted at Camps Sheridan and Taylor under carefully controlled conditions. The results showed that when a person was rated independently by any number of 3 to 13 competent raters, the range in the ratings would commonly be as large as 30 points on a total scale of 80 points. The chances were not more than 4 to 1 that any rating would be within 14 points of the person's true rating.[2]

Rugg believes that it is possible to find raters whose discrimination is accurate and whose judgment of character will correlate very closely with objective measures of it. He quotes the findings of Dr. Chassell who found raters whose judgment correlated 0.7 with objective measures. But she found more whose judgments correlated 0.4 and 0.3 and 0.2 and 0.1 and 0.0 "and the number of such is so large that we dare not use this method of measuring character, with the competency of raters as it exists today."[3]

[1]Harold O. Rugg, "Is the Rating of Human Character Practical?", Jr. of Ed. Psychology, V. XII, No. 8, P 435.

[2]See op. cit. V. XII, No. 9, p 487.

[3]Harold Rugg, "Is the Rating of Human Character Practical?", Jr. of Ed. Psychology, V XIII, No. 2, pp 82, 83.

He continues:

This is the first thing to do, by all means -- increase the number of ratings on a person. Obtain a mass judgment from good judges. . . Assuming qualified raters, the reliability of a judgment increases directly with the square of the number of judgments. To double the reliability, take four times the number of judgments. The Probable Error of a single judgment is 0.67450; of two judgments it is 0.470; of three judgments 0.380; of four judgments 0.340.[1]

THE RELIABILITY OF A SINGLE FACTOR AND SEVERAL FACTORS

Just as mass judgment of competent raters more nearly approaches the truth, so psychologists attempt to secure more accurate results by the measurement of many separate traits. In accord with popular opinion, a teacher may be above average in intelligence test score and in scholarship and yet be a failure as a teacher and vice versa. From all appearances, the teacher may have a good personality and yet be unpopular with the principal and with the class. Ordinarily, it is agreed that a person who is high in one trait is correspondingly high in all other traits. Because this theory does not work out practically, personality is divided into its many elements and these are compared with one another and with different measurements to find the relation between them. Such findings are interesting if not always reliable as shown by the following investigation.

Thayer of Ohio State University in a study of the three types of teacher rating plans quotes Knight on the score card type.

[1]Harold Rugg, "Is the Rating of Human Character Practical?" Jr. of Ed. Psyc., XIII, No. 2, pp 82, 83.

Knight made a study of the correlation between a rater's general estimate of a teacher and his scoring of this same teacher on a specific trait. He had the superintendents, principals, and supervisors of 129 New York teachers rate these teachers on Boyce's score card. Eight of the correlations arrived at are as follows:

1.General teaching ability with general intellectual ability	$\pm$.677,$\pm$.03
2.General teaching ability ability with skill in discipline	$\pm$.787,$\pm$.02
3.General teaching ability with voice	$\pm$.632,$\pm$.04
4.General intellectual ability with voice	$\pm$.625,$\pm$.04
5.General intellectual ability with skill in discipline	$\pm$.560,$\pm$.04
6.Voice with interest in community	$\pm$.500,$\pm$.04
7.Voice with skill in discipline	$\pm$.438,$\pm$.06
8.Skill in discipline with morals	$\pm$.333,$\pm$.11

As Knight states, "Common sense would tell us that the correlation between voice, defined on the score card as 'voice-pitch, quality, clearness of school-room voice'--and interest in community is probably zero, but here it was found to be $\pm$.500, while voice and discipline was $\pm$.438 and general intellectual capacity and voice was .625. The sizes of the correlations do not correspond to the importance of the relationships. In other words, a judge has a certain opinion of a teacher <u>in toto</u>, and his opinion is given according to his general impression in answer to any significant question about that teacher. It seems fair to conclude, that in judging particular traits general estimate influences the particular estimate to such a degree that judgments of particular traits are in themselves of little practical use."[1,2]

After a careful study of the practical applications of such scales as Elliott's, Beatty's, Boyce's, and Hill's, Rugg concludes that ordinary scales should be discarded. He states that "the unreliability of current typical ratings of teachers by principals is so great that it is almost valueless". For example, when the ratings of a large number

[1]V. T. Thayer, "Teacher Rating in the Secondary Schools", <u>Educa. Admin. and Superv.</u>, V. XII, No. 6, pp 366, 368.

[2]See also Knight, F.B., "Qualities Related to Success in Teaching", <u>Teachers College Contribution to Ed.</u>, No. 120, Ch V, N. Y. Teachers College, Columbia Univ., 1922.

of teachers are plotted the curve is badly skewed. In a study of the ratings given to 7131 teachers he found that "96 per cent had been rated either superior, excellent, or good".[1,2]

THE LIMITATIONS OF RATING SCALES

There are various reasons for the use of teachers' rating scales. Perhaps the most important is for use in large school systems for promotion of teachers and for salary increases. However, since this is such an important responsibility, it does seem that each individual case should be considered separately so as to do justice to the teacher and not cause ill feeling among other members of the staff. Another reason is for supervision and for conference. This may be practical for prospective teachers but it is usually discontinued during regular teaching when it really might accomplish results beneficial to the new teacher. Teacher training institutions and commercial teachers' agencies depend almost wholly upon rating scales for the recommending of new teachers and the transfer of experienced ones. But if these scales make very little differentiation and rate practically all of the teachers above average and superior, such measures have degenerated to mere form and possibly have never progressed toward true scientific analysis.

[1] V. T. Thayer, "Teacher Rating in the Secondary Sch.", Ed.Ad. & Superv. V. XII, No. 6, p 366.

[2] See also Rugg, H. O., "Self-improvement of Teachers through Self-rating", El. Sch. Jr., XX, p 671.

That some sort of record is convenient for all concerned, Ruediger agrees, and suggests the personal file which should contain all good and bad records for the principal, the teacher, or the student. He believes that popular rating scales cause a lot of annoyance, humiliation, self-consciousness and apprehension and they give nothing commensurate in return. They tend to injure rather than improve instruction. He does not believe that rating scales can measure the real worth of teachers as evidenced in the following quotation.

> Another point that is not adequately brought out in rating sheets is that the supreme worth of a teacher is quite often dependent on only one point of real excellence, all the others being mediocre or even less. One of the most stimulating teachers that I had in the normal school was a chaotic instructor, a poor disciplinarian, and he took no interest in athletics, debating or other student activities. His only merit was that he had his own ideas about everything that came up in class or that he could bring in by the heels. He questioned and doubted everything already established or about to be established, and he did this not to be smart, but sincerely. . . We began to think about things and to examine them on all sides before giving them our adherence. On any rating scale that I have ever seen this teacher would have made a sorry showing, yet as I look back, he stands for one of the best influences in all my schooling.
>
> Another teacher that I knew in a high school in which I was teaching also appeared to have just one outstanding point of merit. He had the power to stimulate quiet reflective thought. . .It made no difference whether this man taught mathematics, history or Latin, he got the same reflective response, and he reached freshmen as easily as seniors. Again I know of no rating scale that would have done justice to this man.[1]

Rugg also disproves of subjective rating scales on the basis of

[1] W. C. Ruediger, "Rating Teachers", Sch. and Soc., V. XX, No. 505, pp 263 - 268.

their unreliability as stated in the following quotation.

> We would far better give our energies to the attempt to measure it (human character) objectively, than to make subjective judgments of it on point scales. The point cannot be made too emphatically that we should discard these loose methods of rating once and for all. We cannot justify wasting the time of our school administrators and deluding our teachers with fictitious "ratings" and "marks". Even on one of the so-called "standardized" point rating schemes a single rating has little or no scientific validity.[1]

Hamrin made a study of predicting teaching success for the spring graduating class of 1925 of the State Teachers College at Moorhead, Minnesota. The class consisted of 129 members, all of whom taught during the following school year. The rating sheet listed 54 characteristics on a five-point merit scale. . .In comparing the two mean scores for each of the 88 student teachers who had been rated by two different supervisors, it was found that the scores were identical in only two cases. Of the remaining 86 student teachers, 49 were rated higher by the supervisors at the conclusion of the first term of teaching while the other 37 were rated higher by the supervisors of the second term. This difference was of particular interest since several of the supervisors rating the student teachers at the end of the first term were specialists in some one field, such as music, art, physical education, and industrial arts. It was found that the rating of individual supervisors varied a great deal, some persistently rating higher than others.

[1]Harold Rugg, "Is the Rating of Human Character Practical?", Jr. of Ed. Psyc., V XII, No. 8, p 426.

His findings were as follows:

In comparing the ratings given by the superintendents with the ratings given by the supervisors, it was found that the differences were greater than those between the ratings given by the two groups of supervisors. The superintendents were found to rate higher than the supervisors.

The correlations of various factors to teaching success for this group of teachers were as follows:

a. Army Alpha scores and superintendents' ratings ..... .04±.08
b. " " " " supervisors' " ..... .20±.07
c. " " " " total school marks ..... .39±.06
d. " " " " professional marks ..... .33±.07
e. School marks and superintendents' ratings ..... .05±.06
f. " " " supervisors' " ..... .45±.05
g. " " " professional marks ..... .81±.02

(1) The rating is highly subjective; (2) of the fifty-four characteristics only nine showed more agreement than on the entire list--a shorter scale will prove more accurate than the longer scale; (3) the supervisors rated the student teachers lower than did the superintendents; (4) neither scores on the Army Alpha intelligence test nor school marks were a guide to the success of the teachers as measured by the superintendents' ratings. The relation was greater between the supervisors' ratings and both intelligence scores and school marks than between these measures and the superintendents' ratings; (5) there was a marked tendency on the part of the superintendents to rate teachers high or low on the basis of "personal equipment" while the supervisors stressed "technique of teaching" more than did the superintendents; (6) there was evidence that none of the ratings of the training school-supervisors' ratings, school marks, and intelligence scores - were indicative of success as a teacher as measured by the superintendents' ratings. Some teachers with all these ratings in their favor were marked low by their superintendents and vice versa; (7) there is need of a better understanding between the superintendents in the field and the supervisors in the training school as to what constitutes a good teacher. At the present time the definition of a good teacher appears to be highly personal, subjective, and indefinite.[1]

SUMMARY

The majority of rating scales in general use evaluate personality, general preparation, professional attitudes, teaching

[1] S. A. Hamrin, "A Comparative Study of Ratings of Teachers-in-Training and Teachers-in-Service", The Elementary School Journal, V. XXVIII, No. 1, pp 39 - 44.

qualities, management, results of teaching, and extra-classroom activities, all on a five-point merit scale. When students rate their teachers practically these same qualities are mentioned in their ratings.

There is more evidence to prove the reliability of mass opinion than the reliability of any single rating of an individual. Even under carefully controlled scientific experimentation it is almost impossible to secure ratings of character and personality that are sufficiently accurate to be of any value.

Correlations between an objective measure and a subjective measure are not any more accurate than correlations between two subjective measures. In any subjective measure the "halo" effect is present. Subjective ratings of character and personality tend toward high ratings. Rating scales do not do justice to teachers whose excellence is due to traits not mentioned on any rating scale.

CHAPTER IV

INTELLIGENCE TEST SCORES AS PREDICTIVE OF TEACHING SUCCESS

Extravagant claims have been made for the use and the value of mental tests. Perhaps an explanation from an experienced educator will help to give us the proper perspective toward them. Terman says:

> The purpose of intelligence tests is not to deprive anyone of any educational opportunity from which he is fitted by ability to derive normal profit, but rather to enable us to select the type of curriculum from which a given individual can profit, whether he be bright or dull. . . The great value of the intelligence test lies in the fact that it furnishes data not duplicated from any other source. It gives a new line on the student. More than any other kind of information it tells us what grade of work we have a right to expect. It gives a favorable starting point for investigating the causes of failure. It enables us to discriminate between the intelligent student whose failure is avoidable and the student whose inferior native ability renders him comparatively non-educable.[1]

Terman agrees with Wood that the greatest value of mental tests lies in their possibilities for guidance purposes. Gist suggests a minimum score for students electing teaching, but in this study no discrimination was made on the basis of test score for those students electing teaching. During the five-year period of 1923 to 1927, inclusive, 867 entering freshmen of the College of the Pacific took either the Thorndike Examination or the Thurstone

[1] L. M. Terman, from the Introduction in, Measurements in Higher Education, by Wood, 5.

Psychological Test. Of this number it has been possible to secure complete records of 122 who upon graduation from this college entered the teaching profession. The intelligence test scores for all freshmen are divided on the quintile basis. Table I shows the distribution of the 122 teachers in the quintiles. Chart I shows the uneven distribution which is almost the reverse of a normal probability curve.

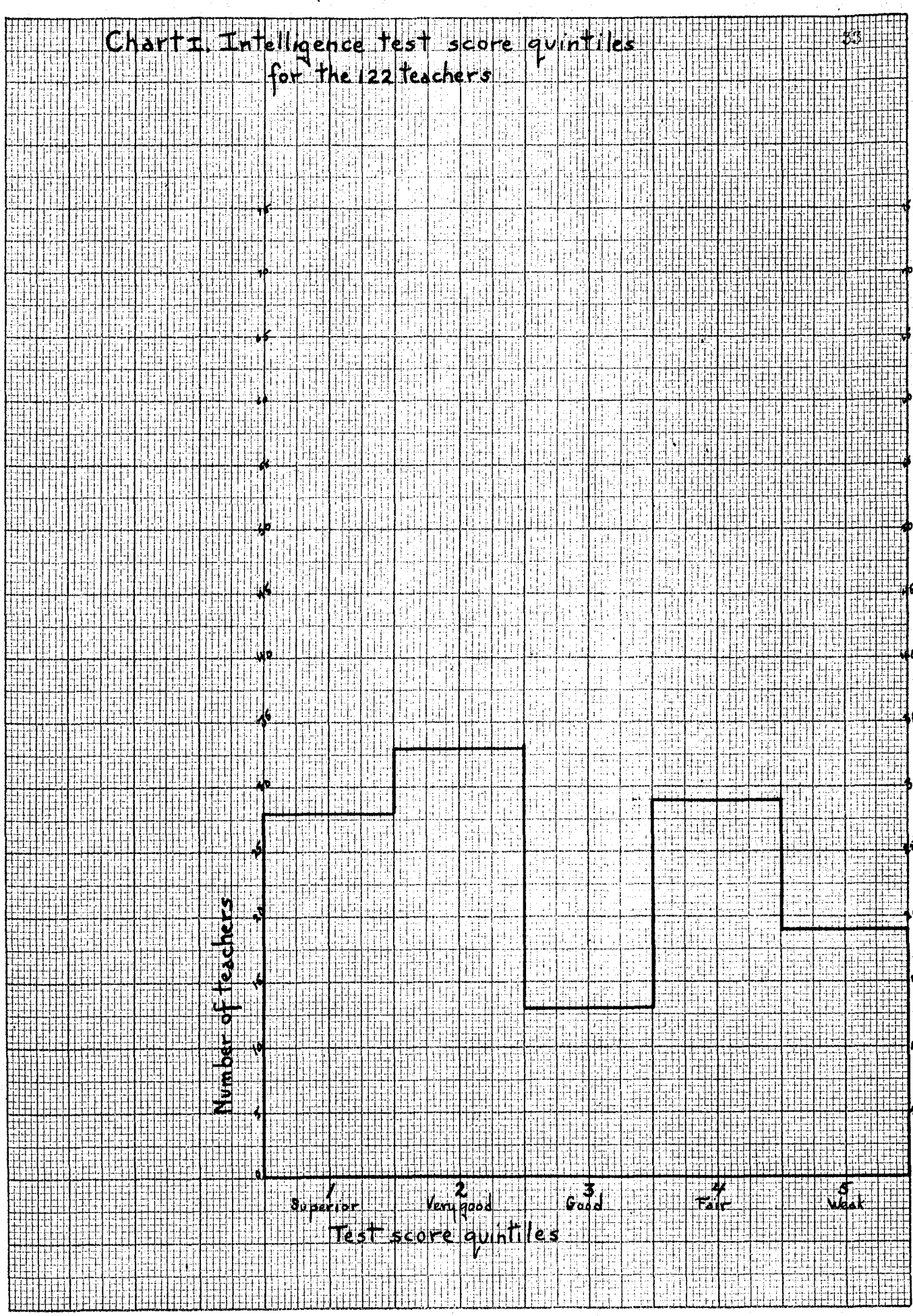
Adjectives are used to interpret the numerical value of the quintiles: namely, 1, superior; 2, very good; 3, good; 4, fair; and 5, weak. On other scales these same values are often stated as follows: excellent, very good, average, fair, and poor. However, in referring to the different steps of the scale such terms as the following will be used in this study; the first quintile, the upper fifth, the lowest quintile, etc.

TABLE I

The distribution of intelligence test scores for the 122 teachers.

Quintile	Number of Teachers	Percentage
1 (Superior)	28	23
2 (Very good)	33	27
3 (Good)	13	11
4 (Fair)	29	24
5 (Weak)	19	15
Total	122	100%

Chart I. Intelligence test score quintiles for the 122 teachers



Analyzing the data presented by Table I and Chart I, the reader notes that there are more teachers in the two highest fifths than in any other two fifths. Less than half the number of teachers in either of the upper quintiles is found in the middle fifth. One low and one high quintile are almost equally represented, namely, the fourth and the first. There is a larger number of teachers in the lowest fifth than in the middle fifth. These teachers are an unusual group as judged by the distribution of their intelligence scores on the quintile basis which is used for all entering freshmen. If teachers are to be a select group on the basis of high intelligence, Chart I should be skewed toward the high ratings with the largest number in quintiles 1, 2, and 3. However, this group of 122 teachers are fairly evenly distributed in both upper and lower quintiles. The relation of this distribution of intelligence test scores with teaching success will now be presented.

Chart II shows the distribution of these 122 teachers according to the intelligence test score quintile and the teaching success group. An explanation of the five-group basis for teaching success is given in Chapter VI. Of the 28 teachers who rank first quintile in intelligence test scores, only 4 of these rank in the first group in teaching success. The others are rated groups 2 and 3 in success. The 33 teachers in the second highest quintile in intelligence test score are rated in group 4 in success. The large majority are on

Chart II The relation of intelligence test score quintiles with teaching success groups for the 122 teachers

Intelligence quintiles	1 Superior	2 Very good	3 Good	4 Fair	5 Weak	
5	4	10	2	3		19
4	6	19	4			29
3	6	4	3			13
2	1	24	7	1		33
1	4	19	5			28
	21	76	21	4		

Teaching success groups

corresponding levels of intelligence quintiles and success groups.

The 13 teachers of the middle intelligence quintile are rated higher in success than in intelligence. Only 3 receive the same rating in both measures whereas 10 are rated higher by their principals. Of the 29 teachers in the fourth intelligence quintile, none are rated below the third group in success. Two thirds of the teachers are rated in the second group and more are rated 1 than 3 in teaching success. While there are no equal ratings in division 5, 3 of the teachers in the lowest fifth of the intelligence test scores are rated in the fourth group, only 2 are in the middle level, and the large majority of the 19 teachers are rated in groups 1 and 2 in teaching success.

SUMMARY

The 122 selected teachers trained at the College of the Pacific from the classes of 1923 to 1927 inclusive are fairly evenly distributed in the quintile divisions of intelligence test scores. Quintiles 3 and 5 have fewer numbers which may indicate that teachers of average and poor intelligence are in the minority. The greatest number is found in quintile 2 which is interpreted in this study as very good, or above the average.

The intelligence test scores are not a reliable criterion for predicting success in teaching. Whereas 48 teachers are grouped in the 2 lowest quintiles on test scores, only 4 teachers are rated in the corresponding groups in teaching success. Likewise, only 61

teachers receive test scores in the 2 highest intelligence quintiles, but 97 are rated in the corresponding groups in teaching success. Only 3 teachers of the 13 who are in the average quintile in test scores are also in the average group in teaching success.

These findings of this study parallel closely the results of other investigations mentioned in Chapters II and III. High intelligence test scores do not predict correspondingly high teaching success when teaching success is measured exclusively on the basis of superintendents' or the principals' ratings of teachers.

Teaching success depends upon factors other than the high degree of general intelligence measured by mental tests.

CHAPTER V

SCHOLARSHIP RECORD AS PREDICTIVE OF SUCCESS IN TEACHING

School marks or grades show some positive relation to success but not a sufficient relation to be used as the sole criterion. This was illustrated by the results of the recent investigations mentioned in Chapters II and III. Before analyzing specific data of the scholarship records of the 122 teachers in this study, it may be well first to consider the meaning and uses of grades as set forth by Symonds in the following quotation.

> (1) to inform pupils and parents of pupils' achievement; (2) as incentives to study; (3) to promote competition; (4) to determine promotion; (5) to determine graduation; (6) to predict a pupil's future success; (7) to enable college authorities to pass on the qualifications of entrance candidates; (8) to determine credits, honors, etc.; and (9) to determine participation in extra-curricular activities.[1]

That the grades ordinarily given to students do not fulfill these various uses is well known. Perhaps there are reasons for the unreliability of school marks. Particularly is this study concerned with number (6) which specifies that marks predict the pupils' future success [whether in school life or post-school life]. Wood enumerates the several bases on which marks are actually given in colleges. They are as follows:

> (1) on effort put forth by the individual student; (2) the general intelligence of the student; (3) the character and personality of the student; (4) the general fitness of the student to live in civilized society; (5) the amount of the improvement in the student

[1] P. M. Symonds, Measurement in Secondary Education, 498.

in general or in specific courses; and (6) the actual achievement of the student in the specific course, in the total school situation, or in the total life situation.[1]

Such variations in marks are inevitable when marks are dependent upon subjective opinions of teachers. Practically all of the experiments on grades are made on just such types of markings. This present study is no exception. Discrepancies in school marks can probably be alleviated by the use of standard and objective tests which measure only the pupils' achievements. Then scholarship may possibly rank with intelligence tests as an objective measure with more or less degree of reliability.

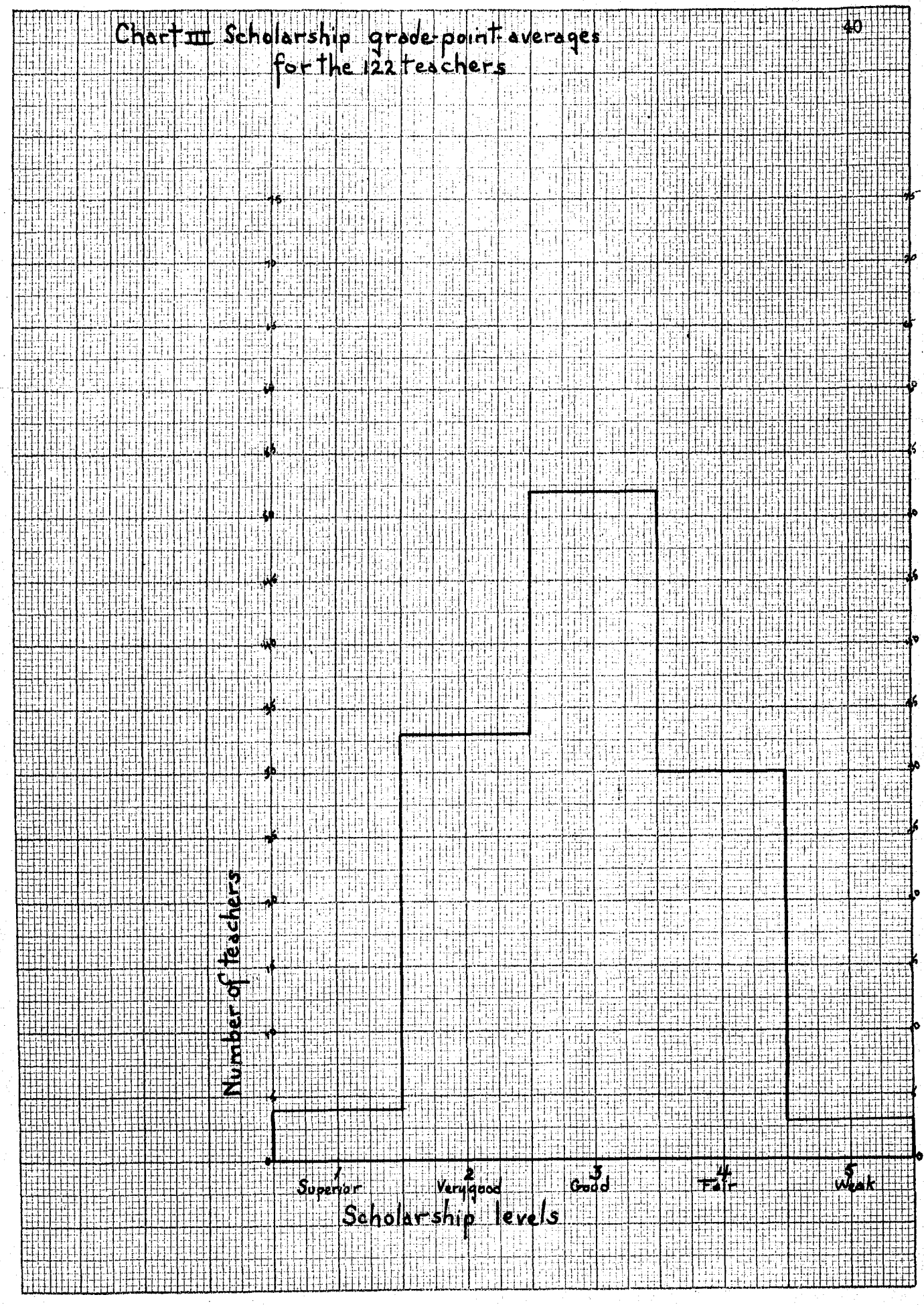
SCHOLARSHIP RECORDS OF 122 TEACHERS

To be graduated from the College of the Pacific the student must have an average of C or a grade-point-average of 1.00. Since all the teachers considered in this study are graduates, the lowest scholarship record is 1.00. For this reason the regular office system for determining the number of honor points is disregarded and a division of measurement more adapted to the study of teachers' scholarships is substituted as follows: level 1 includes all grade-point-averages from 3.00 to 2.50; level 2, 2.49 to 2.00; level 3, 1.99 to 1.50; level 4, 1.49 to 1.01; and level 5, 1.00.

Table 2 shows the distribution in undergraduate scholarship averages for the 122 teachers on this special basis. ChartIII also shows the distribution of these averages. In a large group the normal distribution on a five-point scale may be as follows: 4%, 24%, 44%, 24%, and 4%, respectively. These teachers' averages very nearly approach this

[1]Ben Wood, Measurement in Higher Education, 114

Chart III Scholarship grade-point averages for the 122 teachers



curve and are representative of rather reliable grades given to the graduates of the college. However, if teachers are to be selected on the basis of high scholarship, the distribution will be skewed to the higher averages, even on such a rating scheme as used in this study.

TABLE 2

The scholarship averages of the 122 teachers on the basis of undergraduate college study.

Levels	Grade Point Average	No. of Teachers	Percentage
1 (Superior)	3.00 - 2.50	4	3
2 (Very good)	2.49 - 2.00	33	27
3 (Good)	1.99 - 1.50	52	43
4 (Fair)	1.49 - 1.01	30	25
5 (Weak)	1.00 -	3	2
Total		122	100%

Table 3 shows the selections of college major subjects for the 122 teachers and the range of each subject in scholarship levels. The one student in industrial education is included in education majors and the one student in ancient language is included in the modern language group. The few in dramatic art are included in the speech majors. The number of teachers in each group varies from only one in philosophy to thirty in music. Of the 12 major-groups of subjects only 4 are represented in the highest level of scholarship: namely, English, history, language and mathematics. Only 3 major-groups are represented in the lowest level: namely, education,

music, and speech. Art majors are distributed in levels 2 to 4; economics and social science, 3 and 4; education, 2 to 5; English 1 to 3; history, 1 to 3; languages, 1 to 4; mathematics, 1 to 4; music, 2 to 5; philosophy, 4; physical education, 2 to 4; speech, 2 to 5; and sciences, 2 to 4. English and history major students have the highest scholarship averages, all being rated in the 3 upper levels. Economics and social science major students rate lowest in that none are in scholarship levels 1 and 2, although none are found in the fifth level.

TABLE 3

Scholarship levels in the college major subjects for the 122 teachers.

Level	Art	Econ. & Soc. Sci.	Educ. & Indus. Ed.	Eng.	Hist.	Language	Math.	Music	Philosophy	Phys. Ed.	Speech	Sciences	Total stu. in ea. group
1				1	1	1	1						4
2	2		2	3	6	3	1	8		1	6	1	33
3	3	4	6	4	4	5	1	13		6	3	3	42
4	2	3	4			3	1	8	1	4	1	3	30
5			1					1			1		3
Total	7	7	13	8	11	12	4	30	1	11	11	7	122

TABLE 4

Intelligence quintiles in the college major subjects for the 122 teachers.

Quintile	Art	Econ. & Soc. Sci.	Educa. & Indus. Ed.	Eng.	Hist.	Lang. Anct. & Mod.	Math.	Music	Philos.	Phys.Ed.	Speech & Dr: Art	Sciences	Total in ea. quintile
1	1	2	4	3	1		3	7	1	1	2	3	28
2	1	2		1	5	6		11		3	3	1	33
3	2	1	2	1	1	2		2			1	1	13
4	1	1	4	2	4	2		9		1	4	1	29
5	2	1	3	1		2	1	1		6	1	1	19
Total	7	7	13	8	11	12	4	30	1	11	11	7	122

MAJOR SUBJECTS AND INTELLIGENCE QUINTILES

Table 4 is similar to Table 3 and shows the distribution of college major subjects in the intelligence-score quintiles. Students often consider one subject more difficult than another and do not hope to receive high scholarship while other subjects, particularly those of content (in contrast to scientific studies), are relatively easy and rather high grades are inevitable. This theory was illustrated by data of Table 3 but was not fully proved. Another like theory applies to major subjects and degrees of intelligence. In present practices for guidance, a student must receive a certain

test score to become eligible to special types of training. For the average individual with no particular aptitude this plan might be practical. However, experiments have shown that standard intelligence tests do not measure special talent or specific traits which make for success for certain individual cases. Therefore, no one subject attracts all the bright students, nor does another attract the dullards, but all levels of intelligence are represented in each subject. While it is true that there are general tendencies in one subject or another to require high intelligence, it is not true that a student would necessarily have to fail if he had a lower level of intelligence.

In Table 4, 6 of the 12 major subjects are represented in each of the 5 quintiles in intelligence; 4 major subjects are in the 4 quintiles; and 2 are represented in a smaller range of quintiles because of the few students in each. Art has greater numbers in intelligence quintiles 3 and 5; economics, 1 and 2; education, 1 and 4; English, 1 and 4; history, 2 and 4; languages, 2; mathematics, 1; music, 2 and 4; philosophy, 1; physical education, 2 and 5; speech, 2 and 4; and sciences, 1. Thus, one subject which attracts students of high intelligence is also noted to contain almost as many other students in the lower intelligence quintiles. However, if an approximate rating is desired, all the students in the 3 upper quintiles of the intelligence test scores could be grouped in rank order of their major subjects as follows; the highest, philosophy (only 1 student represented), then mathematics, languages, music, economics,

sciences, art, English, history, speech, education, and the lowest, physical education.

When the 12 major subjects are arranged according to high scholarship, there is a decided change from the above order arranged for intelligence. For example, history and English ranked ninth and eighth in intelligence order but first and second in scholarship order. Likewise, economics ranked fifth in intelligence order but twelfth in scholarship.

TYPES OF CREDENTIALS

Table 5 shows the distribution of the types of credentials granted to the 122 selected teachers graduated from the College of the Pacific. The special credentials include those teachers particularly qualified to teach in music, art, physical education, and speech. This table lists only 1 credential for each teacher, the highest credential selected if there are 2 or more. Twenty-eight teachers of this group have received more than one credential either granted upon graduation or after post graduate work.

TABLE 5

Credentials for the 122 teachers.

Type	Frequency
Junior High	12
Elementary	20
Special	36
Secondary	54
Total	122

SCHOLARSHIP LEVELS AND TEACHING SUCCESS

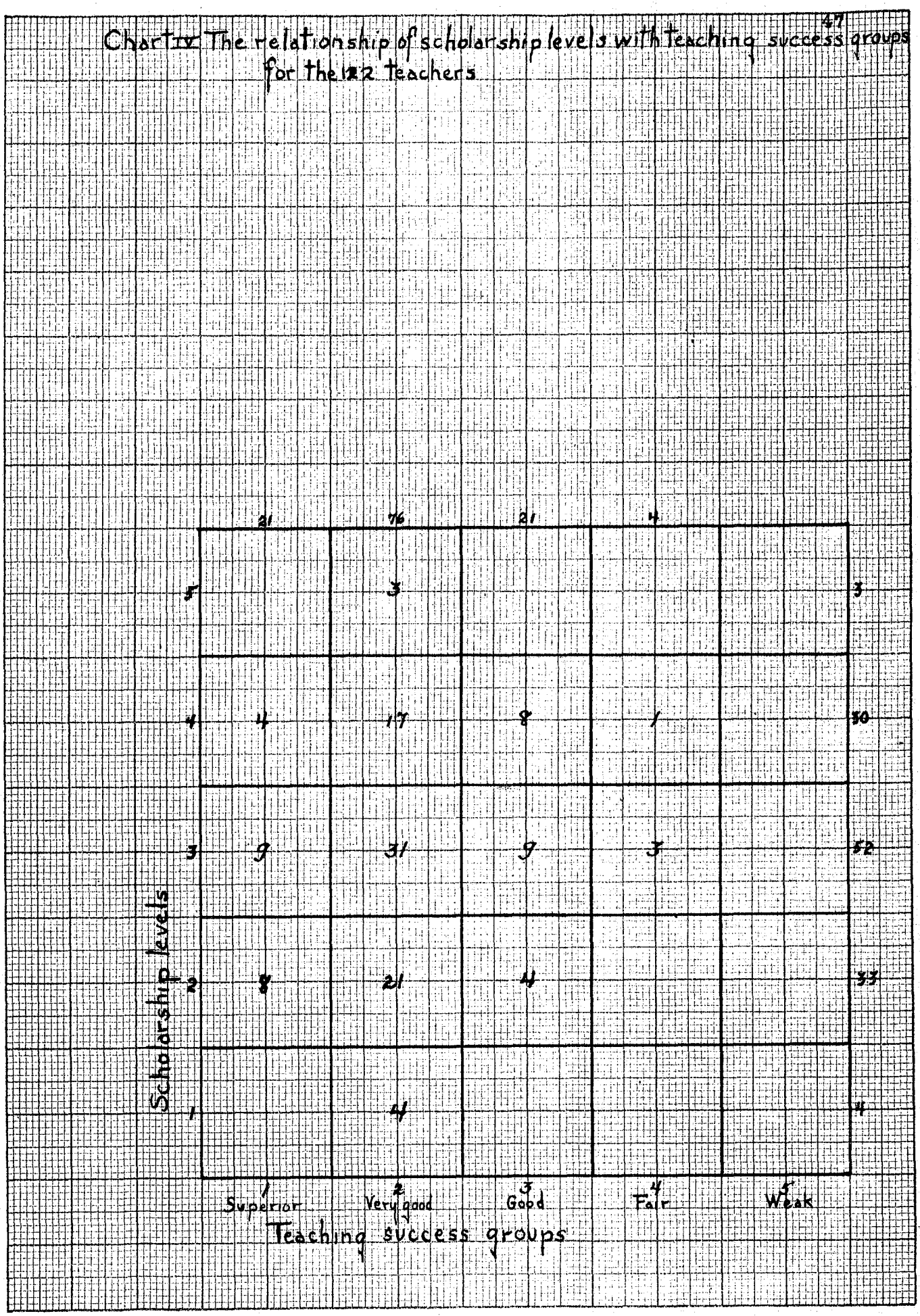
Chart IV shows the relation of scholarship with success. The 4 teachers in level 1 in scholarship all rated group 2 in success. Of the 33 teachers in scholarship level 2, 29 rated groups 1 and 2 in success and 4 rated below. Of the 52 teachers in scholarship level 3, 40 rated group 1, 2, and 3 in success, and only 3 below. Of the 30 teachers in scholarship level 4 only 1 teacher rated success group 4. The 3 teachers in the lowest scholarship level rated in the second highest group in success. The best teachers and the poorest teachers in scholarship were all rated in the same teaching success group. Of the total 122 teachers only 31 were rated in corresponding levels and groups of scholarship and success. With the exception of scholarship level 1 the large majority of the teachers were rated in the success groups above their scholarship level.

SUMMARY

The grade-point-averages as expressed in scholarship levels for the 122 selected teachers graduated from the College of the Pacific very nearly approach the distribution of the normal probability curve. Scholarship ranges from the highest to the lowest in a rather exact proportion.

The major subjects of music and education attract the greatest number of teachers whereas philosophy and mathematics attract the least. Teachers whose major subject is history or English rate the highest in scholarship, and teachers majoring in economics rate the lowest.

Chart IV The relationship of scholarship levels with teaching success groups for the 122 teachers



Scholarship records and intelligence test scores are not highly predictive of each other. Major subjects attract students of all levels of intelligence. However, some teachers whose test scores are high elected philosophy, (only one student represented), mathematics and languages; whereas other teachers whose test scores are low elected speech, education, and physical education. There are discrepancies in these ratings because each group is not represented with an equal number of teachers.

The smallest number of the 122 selected teachers were granted the Junior High School Credential whereas the greatest number received the General Secondary Credential.

High scholarship does not predict teaching success nor does low scholarship predict teaching failure. On the contrary this study showed that teachers of the highest scholarship and teachers of the lowest scholarship were rated in the same groups in teaching success. Teaching success is usually rated above the scholarship level, seldom in the same level and rarely below the scholarship level.

CHAPTER VI

RECOMMENDATIONS OF MAJOR PROFESSORS AND CRITIC TEACHERS AS PREDICTIVE OF TEACHING SUCCESS

The present status of the teaching profession is set forth in the following quotation:

> Teaching service is now among the fields of work in which the public insists upon some measure of protection against incompetence. This has not always been true. In our early history the formal instruction of children was considered neither sufficiently complex and difficult nor sufficiently vital to require teachers of high calibre. However, as our philosophy of education has broadened and deepened to include many different aims, services, and technical procedures, and as society in general has gradually realized the importance of the influences under which the child develops, minimum requirements for admission to teaching have been established.[1]

This expansion in the field of education in the range of purposes, abilities, preparation of teachers, and in the larger attendances of pupils makes the problem of selecting teachers and measuring the results of teaching more imperative than formerly. Ellefson found that the trend in secondary education is away "from entrance requirements of large numbers of prescribed courses and to a more liberal choice [of subjects] and to stricter personal standards of capacity, industry and quality of preparation".[2] The stress today is upon a well developed personality rather than upon a high degree of either intelligence or scholarship as is shown by recent investigations mentioned in Chapters II and III, and since society has

[1] "Training Requirements for Newly Appointed Teachers", Research Bulletin of the Nat. Ed. Ass'n., Jan. 1932, p 5.

[2] E. Ellefson, "Predicting Scholarship", Master's Thesis, University of California Library, 1928.

been educated to the importance of the child, the school, the curriculum, and the teacher, "these standards cannot be maintained unless idealistic, practical, and gifted youth is challenged by the privileges and opportunities of educational service and leadership".[1] In an attempt to maintain a standard of efficiency the majority of teachers are selected or rejected on the basis of information obtained from rating scales.

The Teacher Placement Bureau of the College of the Pacific requires that several recommendations from competent persons be given on a form or rating sheet for every registered candidate for a teaching position. These ratings include the specific items as follows: personality, knowledge of subjects, judgment, health and vigor, personal appearance, power of oral expression, energy and persistence, culture and refinement, community interest and standing, ability in discipline, skill as an instructor, and influence on students. A five-point scale for rating is used on the basis of superior, very good, good, fair, and weak.

THE GROUP BASES FOR RATINGS

In many measurements the highest and the lowest groups are the most significant measures. The interpretation of the ratings given by the major professors, critic teachers, and principals or superintendents for the 122 selected teachers is such that there will be a significant relation of one group with another. To faciliate the

[1] O. E. Hertzberg, "New Rigid Entrance Requirements for Teacher-Training Institutions", Sch. and Soc., Nov. 21, 1931, p 702.

use of comparisons between the various factors employed in this study, a five-point scale is used throughout. But this scale is not consistent for all factors because of the uneven distribution of numbers of each factor which makes a true quintile basis impossible. Therefore, intelligence test scores are on a basis consistent with all the 122 teachers' test scores; scholarship averages are on a basis consistent with all the 122 teachers' scholarship averages; and finally, ratings by major professors, critic teachers, and principals or superintendents are consistent with all ratings given to the 122 teachers.

MAJOR PROFESSORS' RATINGS

Because of the major professors' closer contacts and longer period of observation of the prospective teachers, their recommendations are selected from all other college teachers' ratings for the selected group of 122 teachers. Possibly, the reliability of these reports can be questioned because, in order to give the prospective teacher every chance to secure a position, the major professors are likely to emphasize the good qualities and omit the mention of the less desireable ones. The ratings for this study include only the first 9 items on which the major professors are competent to judge. In the preceding chapter, Tables 3 and 4 show that there is no particular relation between the selection of a major subject and high intelligence, or between the major subject and high scholarship. Therefore, the test scores and grades would not unduly influence the rating of the major professors for the prospective teachers in their classes.

TABLE 6

The major professors' ratings on the personality traits for 122 teachers.

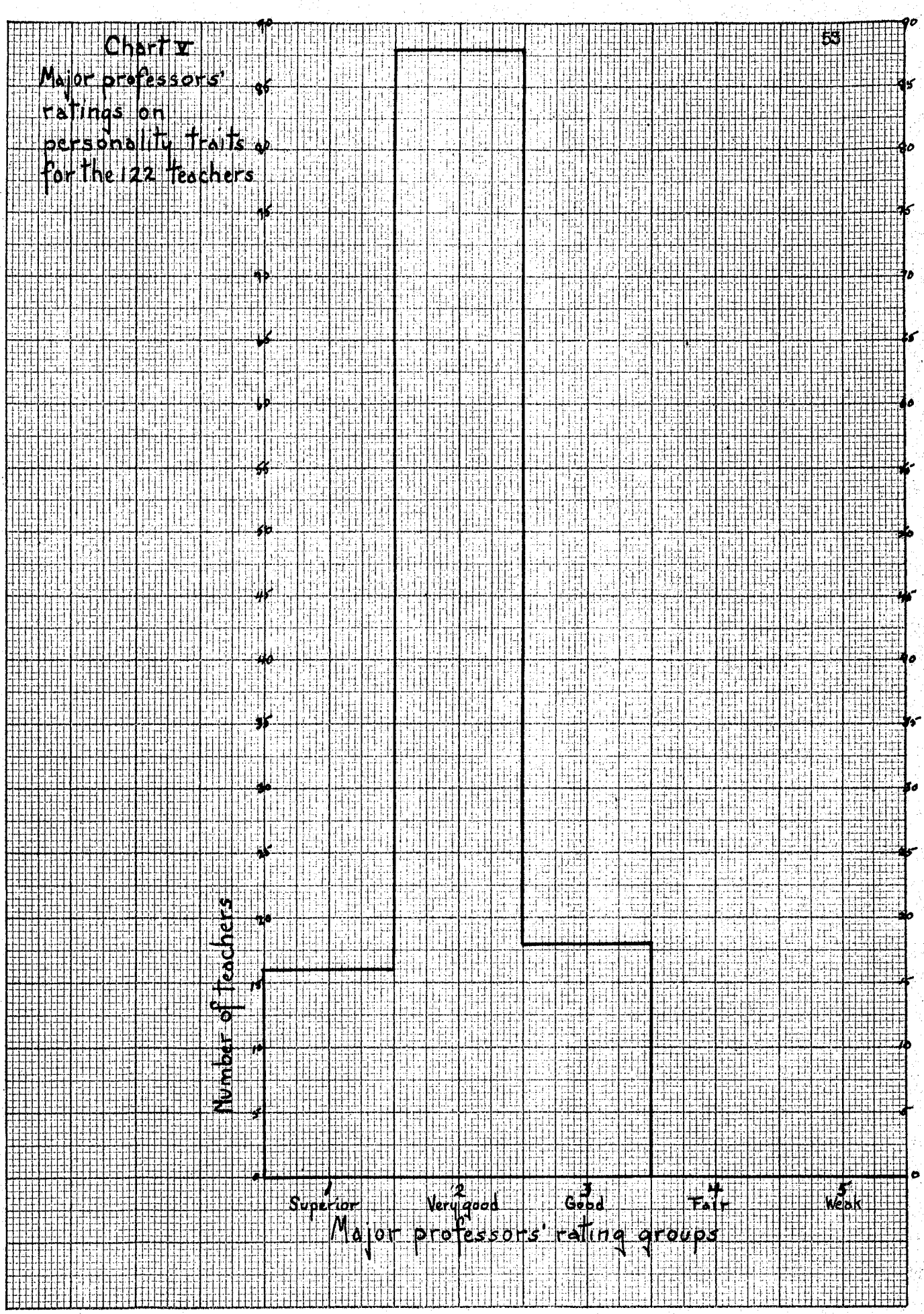
Group	Ratio of Total Estimate	No. of Teachers	Percentage
1 (Superior)	1.00 - 1.50	16	13
2 (Very Good)	1.51 - 2.50	88	72
3 (Good)	2.51 - 3.50	18	15
4 (Fair)	3.51 - 4.50	--	--
5 (Weak)	4.51 - 5.00	--	--
Total		122	100 %

Table 6 shows the distribution of the major professors' ratings on the personal traits of the 122 selected teachers trained at the College of the Pacific. All of these teachers are rated average or above, not even one being represented in either Group 4 or 5. The large majority are rated in Group 2. Chart V further illustrates these same ratings, which indicate that the major professors believe that these prospective teachers possess those personal traits to such a degree that their probable teaching success is assured.

Table 7 shows the relation of the major professors' ratings to the principals' ratings on teaching success. Whereas, the major professors' modestly rated 16 prospective teachers as superior, the principals rated 21 teachers as such. Professors rated almost

Chart V

Major professors' ratings on personality traits for the 122 teachers



two thirds 88, as very good, the principals gave 76 teachers the same rating. The average teachers numbered almost the same, but the principals found 4 fair which were not predicted as such. However, these subjective ratings, while they present the general trend, are not to be accepted as the final criterion of either prediction of success or success in teaching.

TABLE 7

The relation of the total ratings on personal traits with teaching success for 122 teachers.

Group	Number of ratings on teachers		
	Major Professor	Critic Teacher	Principal
1 (Superior)	16	2	21
2 (Very Good)	88	68	76
3 (Good)	18	44	21
4 (Fair)	-	8	4
5 (Weak)	-	-	-
Total	122	122	122

CRITIC TEACHERS' RATINGS

Critic teachers, because of their training and experience, ought to be able to judge fairly well the prospective teachers' professional aptitudes. Their period of observation usually covers every school day for either 1 or 2 semesters. From their total ratings of the prospective teachers only the 3 items concerning professional ability which they were particularly

observant of were selected for this study; namely, ability in discipline, skill as an instructor, and influence on students.

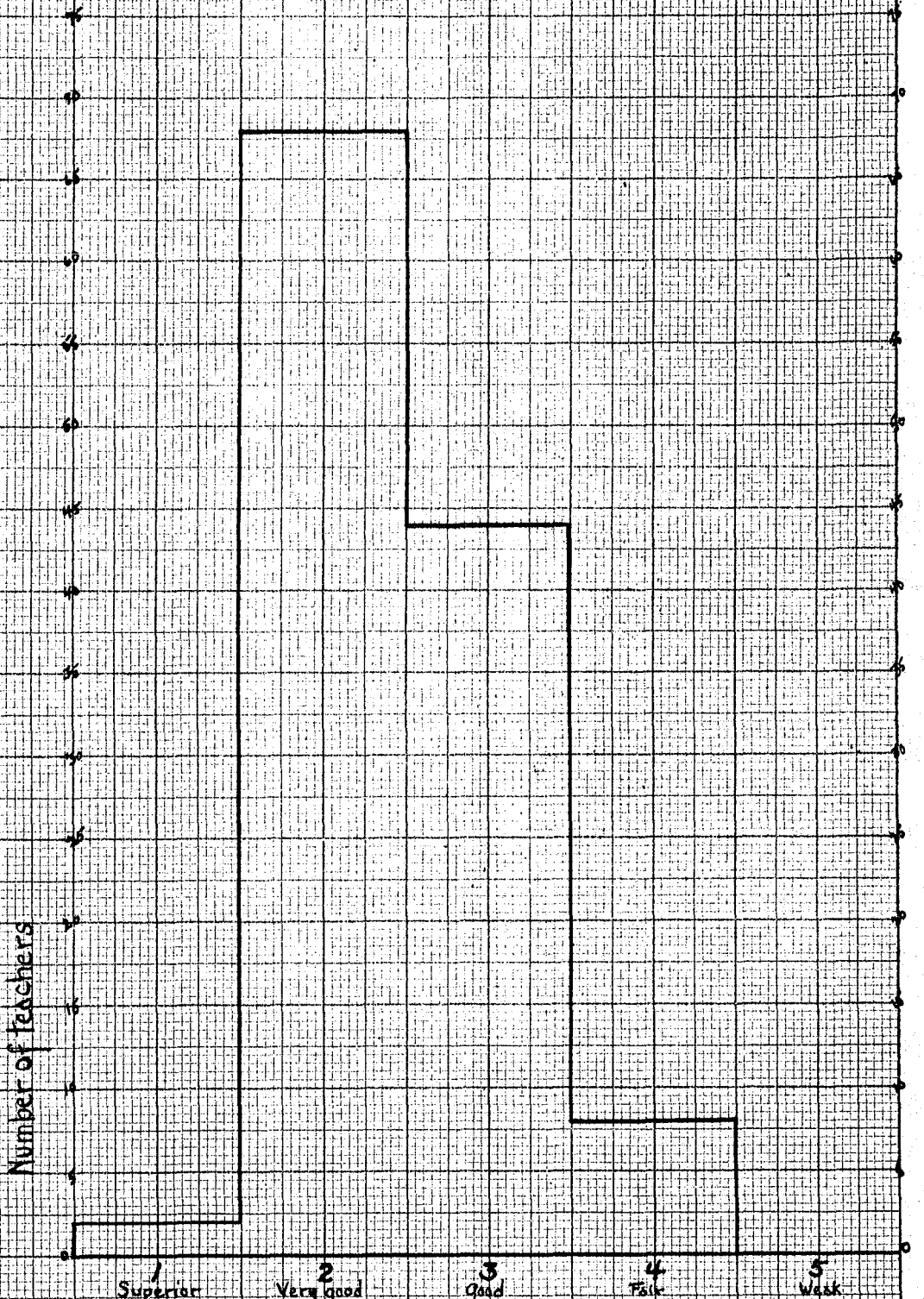
TABLE 8

The critic teachers' ratings on the professional ability of the 122 teachers

Group	Ratio of Total	No. of Students	Percentage
1 (Superior)	1.00 - 1.50	2	2
2 (Very Good)	1.51 - 2.50	68	56
3 (Good)	2.51 - 3.50	44	36
4 (Fair)	3.51 - 4.50	8	6
5 (Weak)	4.51 - 5.00	-	-
Total		122	100

Table 8 and Chart VI show the distribution of the critic teachers' ratings on these 3 items of professional ability. The ratings include the range of only 4 groups. Evidently the prospective teachers do not show themselves as of superior ability in actual classroom situations as shown by only 2 superior ratings. However, more than one half of the total group are rated very good. Critic teachers consider more than one third of the group as average. They, no doubt, have faith in their professional growth and development. Only 8 prospective teachers are rated fair. If the critic teachers are influenced in their ratings by the trait ability in discipline, such ratings will be the most unreliable

Chart XL. Critic teachers' ratings on professional ability for the 122 teachers



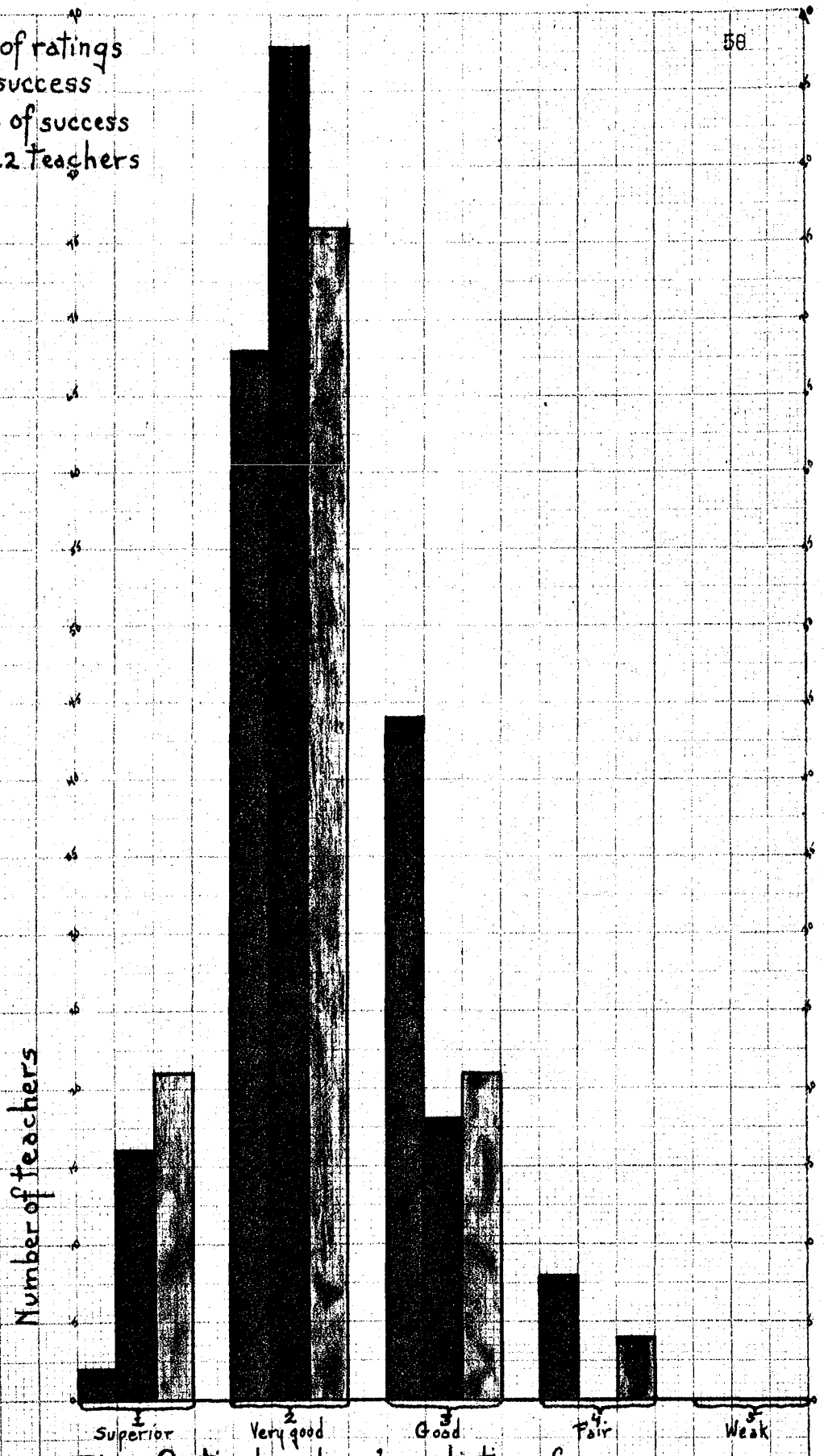
because the student-teacher is not allowed to handle any major problems of discipline. He does not have the opportunity to profit and learn by experience the best methods which apply to difficult classroom situations.

Table 7 compares the critic teachers' partial ratings with the principals' composite ratings. Even though the number of ratings vary, yet the usual disparity is shown between predicted success and teaching success. The critic teachers do not rate student-teachers as high as the principals rate these same teachers.

MAJOR PROFESSORS', CRITIC TEACHERS', AND PRINCIPALS' RATINGS

Chart VII shows the relation of the prediction of success with teaching success. Because the major professors' recommendations represent those ratings on which they are the most competent to judge, and because the same may be said for the critic teachers' ratings, and because the principals must rate on both personality traits and on professional traits, this method of comparing uneven distributions of ratings is apparently justified. There is only one point of exact agreement among the 3 classes of raters. Whereas 1 or 2 specific items may be rated in success Group 5, no total rating given to any teacher by any rater is found in this lowest group in teaching success. The major professors and principals rate the teachers higher than do the critic teachers. All ratings are skewed to the high ratings and

Chart VII. The relation of ratings predicting success with ratings of success for the 122 teachers



Blue-Critic teachers' prediction of success
Black-Major professors' prediction of success
Orange-Principals' ratings of success

all raters apparently agree that the large majority of teachers are in Group 2, very good.

SUMMARY

The same type of rating form is used by major professors and critic teachers to predict teaching success as is used by the principals to measure teaching success for the 122 selected teachers trained at the College of the Pacific. Whereas the major professors and the critic teachers rated all of the items on the rating form, only those items are selected for this study for which the raters give a reliable opinion not based on general impression.

The major professors' ratings are found only in the first 3 groups predicting success. All prospective teachers will be at least average but the large majority will be much above average. The critic teachers' ratings are found in the first 4 of the 5 groups in success. They predict the smallest number of superior teachers and the largest number of fair teachers. The majority of teachers are rated very good. Principals rate a larger number of teachers superior than either the major professors or critic teachers predict as superior. The majority of ratings by the principals is also found in the second success group. On a five-point scale only 3 points are used by major professors, and 4 points by critic teachers and principals to rate teaching success. The lowest point is not represented in any rating.

Major professors and principals rate approximately the same number of teachers superior as they rate average. All raters

consistently place the large majority of the teachers in the second group of success, which is the group above average on a five-point scale.

CHAPTER VII

PRINCIPALS' RATINGS AS THE BASIS FOR DETERMINING TEACHING SUCCESS

Granted that a single rating of an individual has little or no scientific validity, yet it is necessary to accept a single rating for new teachers or inexperienced teachers. Superintendents and principals are in positions to observe and to evaluate their teachers and they ought to be able, by training and experience, to give an unbiased opinion. However, thus far, individual rating scales of character and personality are not designed to reveal the principles governing the administrators who rate some teachers high because of the factors of acquaintance, altruism, personal pride and advantage, favoritism, or indifference; and others who rate teachers low because of the factors of prejudice, impression, or ignorance. Nor do the scales themselves allow for the beginner's enthusiasm or his lack of experience.

In selecting teachers administrators deem it desireable to select those whose recommendations guarantee a good degree of intelligence, scholarship, personality, and professional ability. But from the representative studies quoted in Chapter III it is found that superintendents and principals allow their general impressions of teachers' personalities to influence all specific ratings on other traits of scholarship, etc. Particularly is this true of teachers who have superior personal traits, whom principals rate superior on every other trait.

Knights' studies are rather complete on principals' ratings of teachers. In most instances high correlations existed between unrelated traits and low correlations between related traits. There are many individual interpretations of teaching success but all administrators seem to be consistent in rating high some particular personal traits of teachers.

PRINCIPALS' RATINGS ON SPECIFIC TRAITS

Several specific ratings will be considered in this chapter and only brief mention will be given to each.

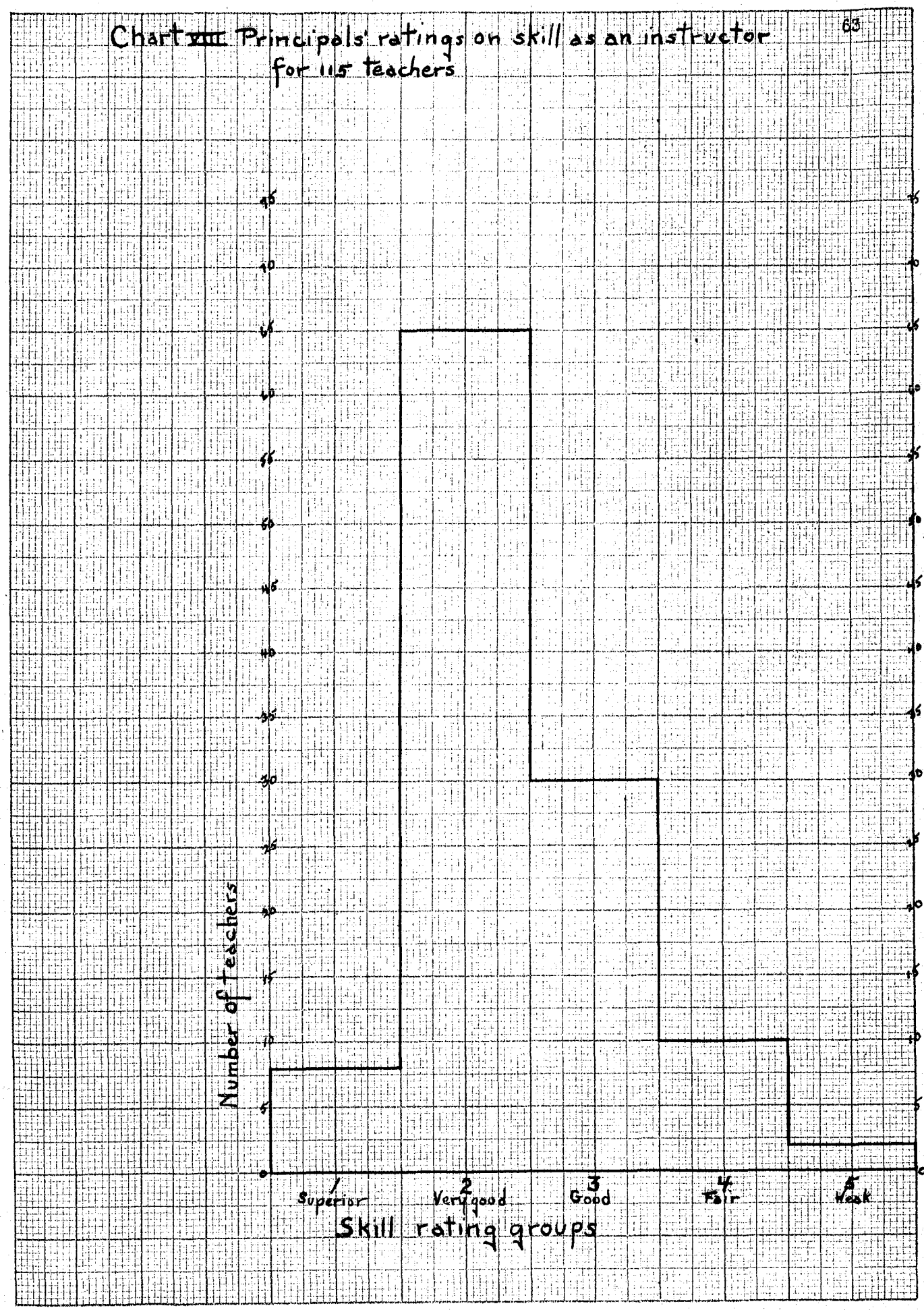
TABLE 9

The principals' specific ratings on skill as an instructor for the 122 teachers.

Group	Number of Teachers	Percentage
1 (Superior)	8	7
2 (Very good)	65	56
3 (Good)	30	26
4 (Fair)	10	9
5 (Weak)	2	2
Total	115	100

Table 9 and Chart VIII show the distribution of the principals' specific ratings on skill as an instructor for 155 teachers. Even on professional items the majority of teachers are rated above average, as shown by the 63 per cent above, the 26 per cent average, and

Chart ~~VIII~~ Principals' ratings on skill as an instructor for 115 teachers



the 11 per cent rated below average. However, these 11 per cent are significant. It is possible that in their preparation for teaching they failed to receive an adequate knowledge of method. It is also true that these teachers may lack innate ability to adapt instruction to the classroom situation.

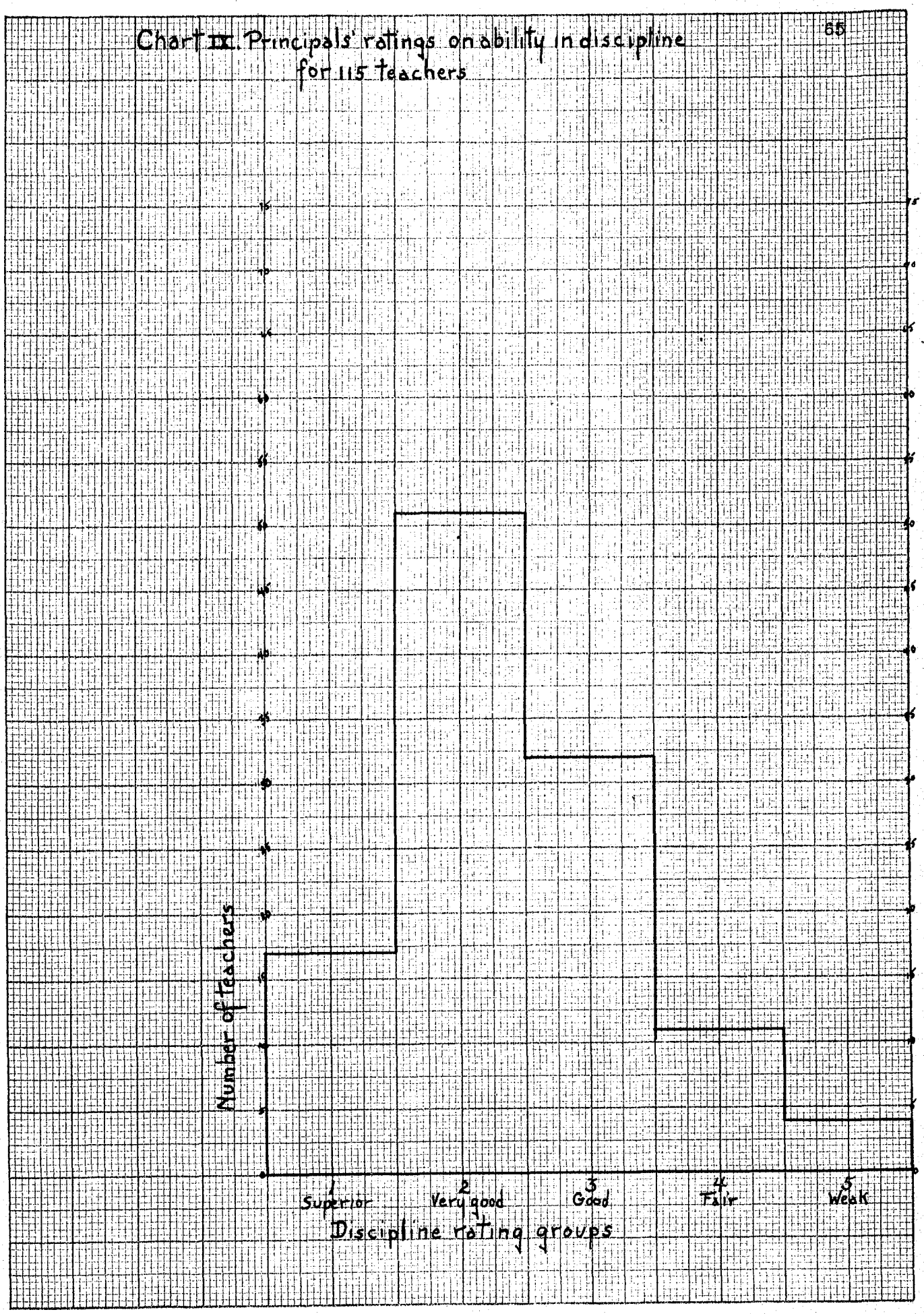
TABLE 10

The principals' specific ratings on ability in discipline for 115 teachers.

Group	Number of Teachers	Percentage
1 (Superior)	17	15
2 (Very good)	51	44
3 (Good)	32	28
4 (Fair)	11	10
5 (Weak)	4	3
Total	115	100

Table 10 and Chart IX show the distribution of the principals' specific ratings on ability in discipline. Although the distribution does not follow the normal curve, yet more variation is shown in the principals' particular rating than in any other rating given by either the major professor or the critic teacher. More teachers are rated superior than fair and weak in ability in discipline. As usual, the large majority are rated very good and good. It is not assumed that all teachers are good disciplinarians. In fact, discipline is usually

Chart IX. Principals' ratings on ability in discipline for 115 teachers



the most difficult problem which confronts the teacher, and in many instances, it is the most important. Until the student-teacher is given more opportunity and authority to handle classroom situations, he must depend upon observation and broad theoretical knowledge for the solution of disciplinary problems. To attain to the ideals of the teaching profession the distribution must be skewed to the higher ratings in each specific rating, such as discipline, as well as in the teacher's general rating.

TABLE 11

Principals' ratings in percentage on three specific traits of 115 teachers.

Group	Subject-matter	Skill	Discipline
1 (Superior)	24	7	15
2 (Very good)	57	56	44
3 (Good)	15	26	28
4 (Fair)	4	9	10
5 (Weak)	-	2	3
Total	100	100	100

Table 11 shows the relation of 3 specific ratings for the 115 teachers whose principals gave detailed ratings. Knowledge of subject matter is rated higher than either skill or discipline.

Although teachers seldom teach subjects equal to their own scholarship levels, it is interesting to notice that only 4 per cent were rated below average in knowledge of subject matter whereas 27 per cent were rated below average in college scholarship.

TABLE 12

The relation in percentage of three specific traits for 1500 new teachers in California in 1930 - 1931.[1]

Quartiles	Subject -matter in per cent	Skill in per cent	Discipline in per cent
Excellent	33	23	25
Good	54	55	49
Fair	12	20	21
Poor	1	2	5
Total	100	100	100

Table 12 shows the relation of these same 3 factors in a study of 1500 new teachers in California in 1930 and 1931. This is represented on a quartile basis so an exact comparison with the teachers trained at the College of the Pacific is impossible. However, the general tendencies are the same. Teachers are rated highest in knowledge of subject matter and lowest in ability in discipline.

[1]Evelyn Clement, "An Evaluation of Teacher-Training", Educational Administration and Supervision, Feb. 1932, p 92.

TABLE 13

Principals' composite ratings on teaching success for the 122 teachers.

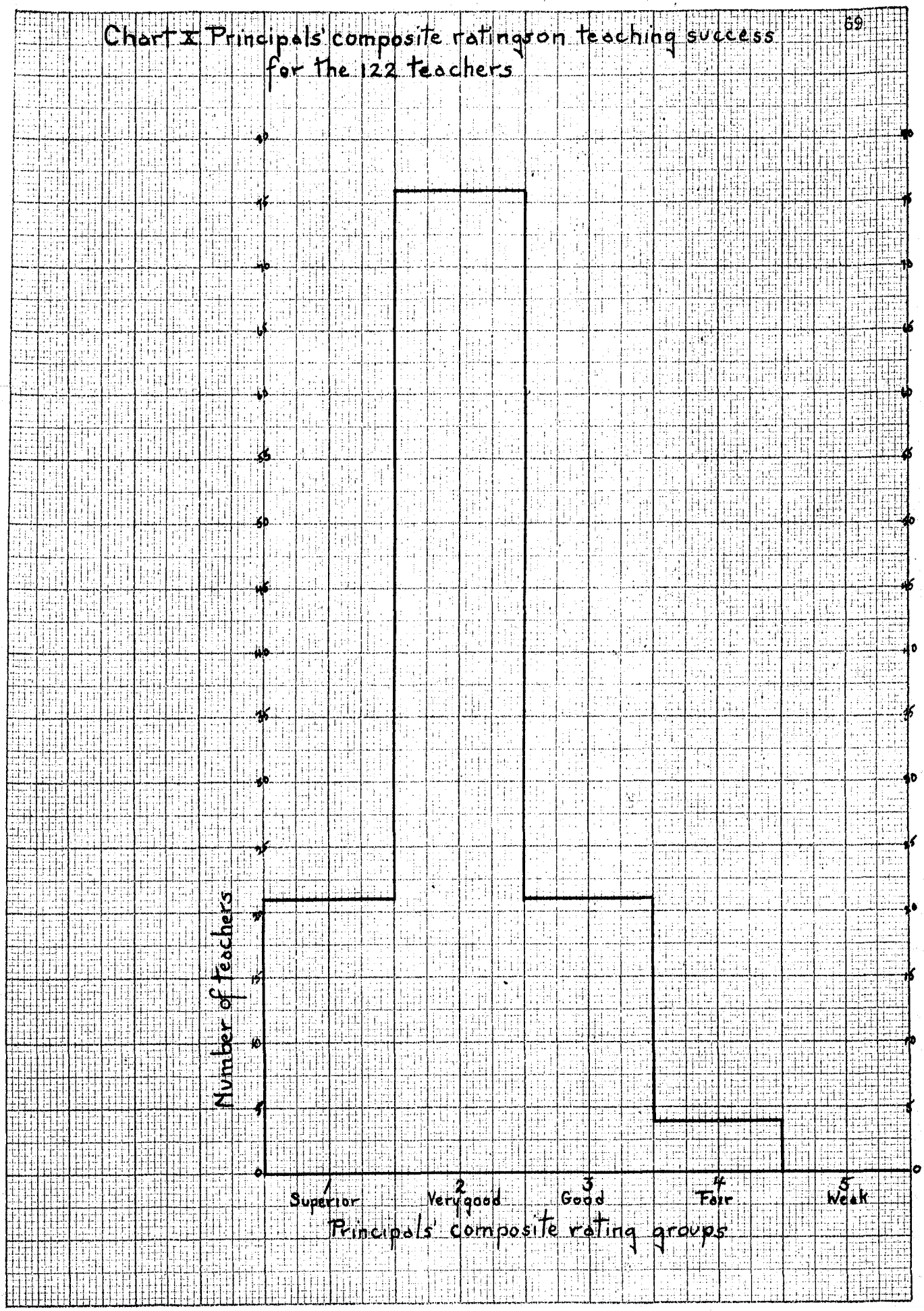
Group	Ratio of Total Rating	No. of Teachers	Percentage
1 (Superior)	1.00 - 1.50	21	17
2 (Very good)	1.51 - 2.50	76	63
3 (Good)	2.51 - 3.50	21	17
4 (Fair)	3.51 - 4.50	4	3
5 (Weak)	4.51 - 5.00	-	-
Total		122	100

Table 13, showing the total ratings, has been used as a basis of comparison in Chapters IV, V, and VI; hence detailed analysis of this table is not required in this chapter. Chart X illustrates the table by showing a rather balanced curve for Groups 1, 2, and 3, but disproportionate numbers for Groups 4 and 5. Almost two thirds of the teachers are rated very good, and the other third is rated equally between the superior and the good groups of principals' ratings.

SUMMARY

Principals' and superintendents' ratings on specific professional traits are lower than their ratings on personality traits

Chart I Principals' composite rating on teaching success for the 122 teachers



for teachers trained at the College of the Pacific. The specific trait of ability in discipline received the lowest rating of the specific traits which are considered in this study. The specific ratings for skill as an instructor and ability in discipline are almost equally divided in the 5 separate groups with the exception of Group 1, which has twice as many superior teachers in discipline as it has in skill. The 2 lower groups contain 13 per cent of the specific professional traits, whereas only 1 lower group contains 3 per cent of the total ratings which include both professional and personal traits.

Low ratings on professional traits are counter-balanced by the higher ratings on personality traits of the total rating scale of 12 items. On such a scheme of rating as is used in this study, principals rate 97 per cent of the teachers superior, very good, and good, and only 3 per cent fair. No teacher is given the general rating of weak or poor.

CHAPTER VIII

GENERAL SUMMARY AND CONCLUSIONS

The purpose of this present study was to examine the various office records of a selected number of teachers trained at the College of the Pacific and to determine the basis for their teaching success. According to other popular methods, various factors are segregated and these are compared with the reports on teaching success. The factors chosen for these selected teachers were the same that are usually considered in similar investigations: namely, intelligence test scores, college scholarship, personality ratings and professional ratings of prospective teachers. These are generally considered to have more or less predictive value of probable teaching success. Of the many types of rating scales for the rating of character, personality, and professional ability, most teacher-training institutions, principals and placement agencies use a scale with an average of about 7 main traits on a five-point merit scale for the rating of success for individual teachers. Such a scale was used by superintendents and principals to rate the 122 selected teachers and these ratings were used as the basis of comparison between predicted success and actual teaching success. The results of this study bear a close relationship with other studies which show in a general way the influence of various factors on teaching success.

This study aimed to find the elements or qualities of teaching success. The ideal is usually much higher than the actual situation. Investigations show that the vast majority of teachers are rated average and above average, regardless of their training and experience or their lack of it, and regardless of the levels of intelligence or scholarship. Should a uniform standard be adopted as the basis of rating teachers, it is probable that there would be a higher degree of correlation between various factors and teaching success. It is also possible that objective rating scales would replace the subjective methods of rating teaching success. Discussion of the relation of tests with teaching success will follow.

INTELLIGENCE TESTS

Neither the Thorndike Examinations nor the Thurstone Psychological Test given to all entering freshmen of the College of the Pacific from 1923 to 1927 inclusive, proves of any predictive value for the teaching success of 122 selected teachers. Of the 97 teachers rated superior and very good in teaching success only 61 received test scores which indicated the same groups. And of the 48 teachers in intelligence quintiles 4 and 5 only 3 teachers are rated in success groups fair and weak. Only 3 teachers are rated good in both intelligence and success.

Investigations have been conducted to find reasons for these inconsistencies. The results of these studies show that intelligence tests do not test the special abilities of a large minority

of successful individuals; that there are many illusive factors which must be recognized before the test score is given reliability; that in the consideration of all other available criteria very high test scores do not predict success any better than very low scores because of the necessary adjustments the genius or the dullard must make to the classroom situation; that some subjects in college and for teaching do require a particular type and degree of intelligence; and finally, that the elements of perserverance, initiative, interest, attitude, etc., are of more value than intelligence test scores.

TABLE 14

The relation in percentage of the principals' composite ratings with intelligence test scores for the 122 teachers.

Group	Principals' ratings	Intelligence Score
1 (Superior)	17	23
2 (Very good)	63	27
3 (Good)	17	11
4 (Fair)	3	24
5 (Weak)	-	15
Total	100	100

Table 14 shows the lack of agreement between the intelligence test scores on a quintile basis and the principals' ratings in teaching success on a five-group basis for the 122 selected

teachers trained at the College of the Pacific.

COLLEGE GRADE-POINT-AVERAGES

The grades received by the 122 selected teachers approach the distribution of the normal probability curve. The major subjects of music and education attract the greatest number of students, and philosophy and mathematics attract the least. Teachers whose major subjects are history or English rate the highest in scholarship levels, and those whose majors are economics and social science rate the lowest. Teachers whose major subjects are mathematics or languages receive the highest intelligence test scores and teachers whose major subjects are education or physical education receive the lowest test scores.

College grades are probably not more dependable in predicting teaching success because professors do not mark consistently on the same factor or factors. Most professors give too many high ratings of A and B which combine personality, industry, accuracy, etc. with scholarship.

TABLE 15

Percentage distribution of principals' composite ratings and scholarship levels for 122 teachers.

Group	Principals' rating	Scholarship
1 (Superior)	17	3
2 (Very good)	63	27
3 (Good)	17	43
4 (Fair)	3	25
5 (Weak)	-	2
Total	100	100

Table 15 shows the lack of agreement between the scholarship levels and the principals' ratings on success for the 122 selected teachers. Regardless of scholarship level more than half, almost two thirds, are rated very good in teaching success. Just as many teachers are rated superior, as are rated good in teaching success. The teachers who rank lowest in scholarship are rated very good, the same rating given those teachers who rank highest in scholarship.

TABLE 16

Percentage distribution of intelligence test scores and scholarship levels for for the 122 teachers.

Group	Intelligence	Scholarship
1 (Superior)	23	3
2 (Very good)	27	27
3 (Good)	11	43
4 (Fair)	24	25
5 (Weak)	15	2
Total	100	100

Table 16 shows the relation between test scores and scholarship. Quintiles 2 and 4 of intelligence score are almost identical with levels 2 and 4 of scholarship, which rate very good and fair, respectively. There is a decided negative relation between quintiles 1, 3, and 5 and levels 1, 3, and 5. These intelligence test scores do not predict probable scholarship for the majority of the 122 teachers.

RECOMMENDATIONS

Certain character, personality, and professional traits are agreed upon as essential to teaching success. Nearly all rating scales of teachers wherever used, contain these indispensable traits.

Only the major professors' ratings on character and personality traits have been selected for this study. Every teacher is given a total recommendation of good, or above. Almost three fourths are rated very good. Whether the ratings are high in order to allow for placement of the teachers or whether personal traits alone are predictive of teaching success is not indicated in the reports. However, the major professors' ratings agree with other subjective ratings in that high ratings are the rule and not the exception. Table 19 shows that even though principals' total ratings are ordinarily high, yet they are only a small per cent lower than the prediction of success by major professors.

TABLE 17

The relation in percentage of principals' composite ratings and major professors' ratings on personality for the 122 teachers.

Group	Principals' Total	Major Professor
1 (Superior)	17	13
2 (Very good)	63	72
3 (Good)	17	15
4 (Fair)	3	-
5 (Weak)	-	-
Total	100	100

Critic teachers' ratings on prospective teachers agree more closely with intelligence test scores and with school marks than do principals' ratings on these same factors. Critic teachers may be considered to give more accurate ratings because they are in daily contact with the teacher and under more varying situations than those afforded to the ordinary principal.

TABLE 18

The relation in percentage of principals' specific ratings and critic teachers' ratings on professional ability of the 122 teachers.

Group	Principals' rating		Cadet Superv. rating
	Skill	Discipline	
1 (Superior)	7	15	2
2 (Very good)	56	44	56
3 (Good)	26	28	36
4 (Fair)	9	10	6
5 (Weak)	2	3	-
Total	100	100	100

Only the 3 professional ratings are selected from the critic teachers' reports on these 122 selected teachers. Table 20 shows the group distribution of the 3 ratings in relation with the principals' specific ratings on 2 of these same 3. Whereas the principals' ratings are higher, they are also lower. Principals' ratings on skill as an instructor and ability in discipline are lower than their other ratings on character and personality traits.

PRINCIPALS' RATINGS

If it is agreed that "even on one of the so-called standardized point rating schemes a single rating has little or no scientific validity," certainly no assertion of reliability can be made for the popular method of rating employed in this study. Although the methods of measurement differ from similar investigations yet the findings generally agree. In a study of the ratings of 7131 teachers Rugg found that 96 per cent are rated either superior, excellent, or good. In this study of 122 teachers trained at the College of the Pacific 97 per cent are rated superior, very good, or good. There is practically no relation between principals' or superintendents' ratings and intelligence test scores (another study shows the correlation .04 ± .06) or between their ratings and teachers' scholarship (another study shows the correlation .05 ± .06).

TABLE 19

The relation of mental ability, scholarship, and teaching success for the 122 teachers.

Groups	Intell. Score		Scholarship		Prin. Report	
	No. of Teach.	Percent	No. Tea.	Perc.	No. T.	Perc.
1 (Superior)	28	- 23	4	- 3	21	- 17
2 (Very Good)	33	- 27	33	-27	76	- 63
3 (Good)	13	- 11	52	-43	21	- 17
4 (Fair)	29	- 24	30	- 25	4	- 3
5 (Poor)	19	- 15	3	- 2	-	-
Total	122	- 100	122	- 100	122	-100

Table 19 and Chart XI show the relation of intelligence test scores, grade-point-averages, and principals' ratings on teaching success, all computed on a five-group basis. None of these 3 factors show an exact relationship to one another. The factors of intelligence tests and grade-point-averages show a closer relationship to each other than either one does to teaching success. If such a wide divergence can be called a relationship at all, there is more relation between intelligence test scores and teaching success than there is between scholarship and teaching success as success is rated by one principal.

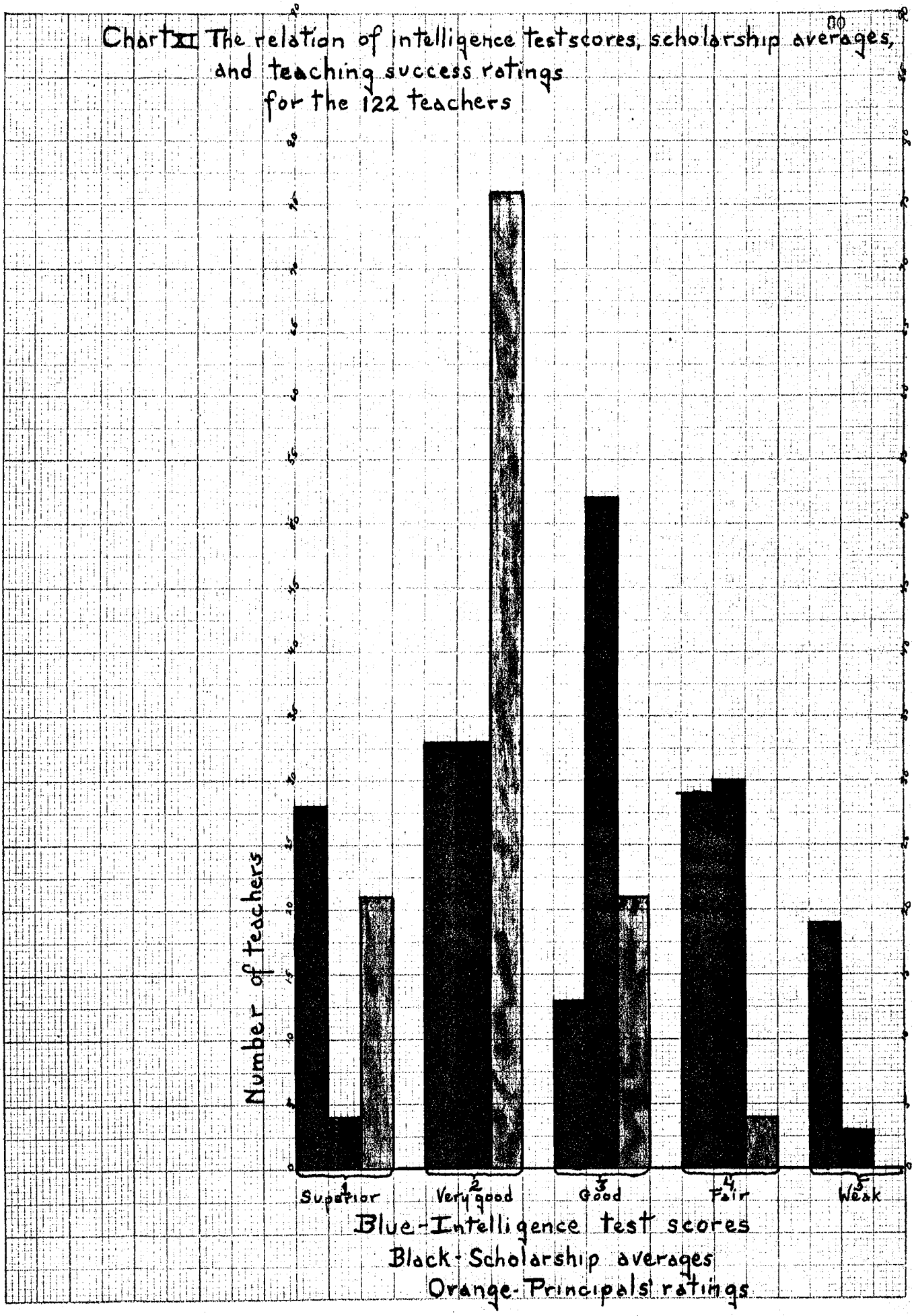
CONCLUSIONS

The results of this study of 122 selected teachers trained at the College of the Pacific in the years 1923 to 1927 inclusive, in regard to mental ability, scholarship and teaching success, are as follows:

Probable teaching success as measured by intelligence test scores and college marks is not predicted with any degree of accuracy, particularly for those prospective teachers with either a very high or a very low intelligence score or a high or a low scholarship record.

Recommendations by major professors on character and personality traits, while uniformly high, do not predict either as many high ratings or as many low ratings for teachers as the principals' ratings show.

Chart XI The relation of intelligence testscores, scholarship averages, and teaching success ratings for the 122 teachers



Blue-Intelligence test scores
Black-Scholarship averages
Orange-Principals' ratings

Critic teachers do not predict many high ratings but predict more accurately the middle and low ratings which agree with the principals' middle and low ratings.

Principals and superintendents do not rate teachers consistently on these traits recognized as indispensable to teaching success. When a discrimination in ratings is made the few professional traits are rated lowest and the many personality traits are rated highest. Thus the composite rating is high.

Teachers rated Group 1, superior, in either intelligence or scholarship are not consistently rated superior in teaching success. The majority are rated very good and good but none are rated either fair or weak.

Teachers in Group 2 of either intelligence or scholarship generally rate the same Group 2, good, in teaching success. However, a larger number of teachers are rated below their intelligence or scholarship group than are rated above them.

Teachers in Group 3 of intelligence or scholarship, generally rate above good in groups very good and superior in teaching success.

Only one teacher in Group 4, fair, of scholarship was rated fair in teaching success. All teachers in intelligence Group 4, fair, are rated good, very good, and superior in teaching success.

Fourteen teachers of the 19 in intelligence Group 5, weak are rated very good and superior in teaching success. The 3 teachers

in scholarship Group 5, weak, are rated either very good or superior in teaching success. No teacher of the 122 selected teachers trained at the College of the Pacific receives a composite rating in Group 5, weak, in teaching success as this success is measured by one superintendent's or principal's rating.

APPENDIX

TABLE 20

The Relation of the Total Individual Ratings for the 122 Teachers.

Case Number[1]	Major Prof.	Critic Teacher	Principal
1	2	2	2
2	2	2	2
3	2	3	2
4	2	3	3
5	2	3	1
6	2	2	2
7	3	3	2
8	2	3	2
9	2	3	2
10	2	2	1
11	2	1	2
12	2	2	2
13	2	4	2
14	3	3	2
15	2	3	2
16	2	2	2
17	1	3	3
18	2	3	3
19	2	3	2
20	2	2	2
21	3	4	4
22	3	2	2
23	3	3	1
24	2	2	2
25	2	2	2
26	2	2	2
27	2	2	1
28	2	2	1
29	1	2	2
30	2	2	2
31	2	2	3
32	2	3	2
33	2	3	4
34	2	2	2
35	3	3	2
36	2	3	2
37	2	2	2
38	2	2	3
39	2	3	1
40	3	2	3
41	2	2	2
42	2	3	3

[1]Case numbers refer to names on file in the office of the Dean of Ed. of the College of the Pacific.

Case Number	Major Prof.	Critic Teacher	Principal
43	2	3	2
44	2	2	2
45	2	3	2
46	3	4	3
47	3	2	2
48	2	3	2
49	2	3	2
50	1	2	2
51	1	2	2
52	2	3	2
53	1	2	2
54	2	2	2
55	3	2	3
56	2	3	3
57	2	2	1
58	2	2	1
59	2	2	2
60	2	2	1
61	2	3	1
62	2	2	1
63	1	2	3
64	1	2	2
65	2	3	3
66	2	4	2
67	2	2	1
68	2	2	2
69	2	4	3
70	1	2	2
71	2	3	1
72	2	2	3
73	3	2	4
74	3	3	2
75	2	3	2
76	2	3	1
77	1	2	2
78	2	2	2
79	2	2	1
80	1	1	2
81	2	3	2
82	2	3	3
83	1	2	2
84	2	2	2
85	2	4	3
86	2	2	3
87	1	3	3
88	2	3	2
89	2	3	2
90	2	3	2

Case Number	Major Prof.	Critic Teacher	Principal
91	3	4	3
92	2	3	2
93	2	3	2
94	2	2	2
95	2	2	1
96	2	2	2
97	2	2	2
98	2	2	2
99	1	2	2
100	2	3	3
101	3	3	2
102	2	4	2
103	3	2	1
104	3	3	4
105	3	2	1
106	2	2	2
107	2	3	2
108	2	2	2
109	2	2	2
110	1	2	1
111	2	3	2
112	2	2	2
113	2	2	2
114	1	3	1
115	2	2	3
116	3	2	2
117	2	2	2
118	2	2	2
119	1	2	2
120	2	2	2
121	2	2	1
122	2	2	2

Explanation

1 Superior
2 Very Good
3 Good
4 Fair
5 Weak

TABLE 21

The relation of the principals' specific ratings with their composite ratings for the 122 teachers.

Case Number[1]	Ability in Discipline	Skill as an Inst.	Total Rating
10	1	1	1.00
28	-	-	1.00
71	1	1	1.00
79	-	-	1.00
105	1	1	1.08
114	1	1	1.08
62	1	1	1.16
60	1	2	1.16
110	1	1	1.16
23	2	2	1.25
61	2	2	1.33
39	1	2	1.33
58	1	1	1.33
27	1	2	1.36
57	1	2	1.41
76	2	2	1.41
121	1	2	1.41
5	2	2	1.50
67	2	2	1.50
95	1	2	1.50
103	2	2	1.50
20	2	2	1.56
77	2	2	1.58
88	2	2	1.58
113	2	2	1.60
81	3	3	1.60
14	2	2	1.66
35	3	2	1.66
44	2	2	1.66
80	1	2	1.66
90	3	2	1.66
109	2	2	1.66
111	3	2	1.66
122	2	2	1.66
6	2	2	1.75
49	2	2	1.75
68	2	2	1.75
94	2	2	1.75
99	2	2	1.75

[1]Case numbers refer to names on file in the office of the Dean of Education of the College of the Pacific.

Case Number	Ability in Discipline	Skill as an Inst.	Total Rating
102	2	2	1.75
117	1	2	1.75
34	2	2	1.80
116	2	2	1.81
15	2	2	1.83
24	2	2	1.83
25	2	2	1.83
118	2	3	1.83
54	1	2	1.83
112	1	2	1.83
7	2	2	1.90
2	3	2	1.91
70	2	2	1.91
51	2	2	1.91
9	2	3	1.91
3	2	2	2.00
11	2	2	2.00
19	-	-	2.00
29	3	2	2.00
30	2	3	2.00
41	-	-	2.00
43	2	2	2.00
47	-	-	2.00
53	2	2	2.00
78	3	2	2.00
84	2	2	2.00
92	2	2	2.00
98	3	2	2.00
119	-	-	2.00
1	3	2	2.08
8	3	2	2.08
12	2	2	2.08
13	2	3	2.08
22	3	3	2.08
37	3	2	2.08
83	2	3	2.08
93	3	2	2.08
101	2	1	2.08
106	2	3	2.08
108	3	2	2.08
59	2	2	2.16
64	2	2	2.16
74	2	2	2.17
16	3	3	2.16
66	3	3	2.25
96	4	3	2.25

Case Number	Ability in Discipline	Skill as an Inst.	Total Rating
36	4	4	2.33
48	3	3	2.33
50	3	3	2.33
52	2	3	2.33
97	3	3	2.33
26	2	3	2.41
32	3	3	2.41
75	3	3	2.41
107	2	3	2.41
45	3	3	2.50
89	2	3	2.50
120	3	2	2.50
31	3	2	2.58
55	3	3	2.58
72	2	2	2.58
82	3	3	2.58
86	4	2	2.58
4	2	2	2.66
85	4	4	2.66
38	3	5	2.83
18	4	4	2.83
91	4	3	2.83
100	5	4	2.83
42	4	3	2.91
65	3	3	2.91
17	3	3	3.00
40	-	-	3.00
63	3	3	3.00
69	3	3	3.00
87	4	4	3.00
115	5	4	3.00
46	4	3	3.16
56	3	4	3.16
104	4	4	3.50
21	5	4	3.83
33	4	4	3.91
73	5	5	4.18

Explanation

1.00 - 1.50	Superior	1
1.51 - 2.50	Very Good	2
2.51 - 3.50	Good	3
3.51 - 4.50	Fair	4
4.51 - 5.00	Weak	5

TABLE 22

The Relation of the Intelligence Score, the College Grade-Point-Average, and the Principals' Composite Rating for the 122 Teachers.

Case Number	Intelligence Score	G.P.A.	Prin. Report
1	2	2	2
2	5	3	2
3	4	3	2
4	2	3	3
5	5	4	1
6	2	3	2
7	4	5	2
8	2	4	2
9	3	3	2
10	5	3	1
11	4	3	2
12	5	3	2
13	4	3	2
14	3	4	2
15	2	3	2
16	2	2	2
17	4	3	3
18	3	3	3
19	4	4	2
20	4	3	2
21	2	4	4
22	5	4	2
23	4	3	1
24	1	4	2
25	2	2	2
26	1	3	2
27	1	4	1
28	1	2	1
29	4	2	2
30	5	4	2
31	4	3	3
32	1	2	2
33	5	3	4
34	2	3	2
35	4	4	2
36	1	3	2
37	2	2	2
38	2	4	3
39	4	4	1
40	5	4	3
41	1	2	2
42	1	3	3

Case Number	Intelligence Score	G.P.A.	Prin. Report
43	1	2	2
44	4	3	2
45	5	3	2
46	2	4	3
47	1	4	2
48	3	3	2
49	1	2	2
50	2	2	2
51	1	2	2
52	4	3	2
53	2	1	2
54	5	4	2
55	1	4	3
56	1	2	3
57	4	3	1
58	3	2	1
59	1	1	2
60	3	3	1
61	3	2	1
62	3	2	1
63	2	2	3
64	1	2	2
65	4	3	3
66	5	3	2
67	4	3	1
68	2	4	2
69	5	4	3
70	2	3	2
71	4	3	1
72	1	2	3
73	5	3	4
74	2	3	2
75	4	5	2
76	1	2	1
77	4	3	2
78	1	1	2
79	3	3	1
80	5	3	2
81	1	2	2
82	4	4	3
83	1	3	2
84	1	2	2
85	2	4	3
86	2	3	3
87	2	3	3
88	2	4	2
89	2	4	2
90	1	4	2

Case Number	Intelligence Score	G.P.A.	Prin. Report
91	3	4	3
92	4	3	2
93	1	3	2
94	1	3	2
95	3	2	1
96	5	4	2
97	4	3	2
98	2	3	2
99	2	1	2
100	3	3	3
101	2	4	2
102	4	4	2
103	5	4	1
104	5	3	4
105	2	3	1
106	4	3	2
107	5	4	2
108	4	2	2
109	2	2	2
110	4	2	1
111	3	3	2
112	2	2	2
113	4	3	2
114	5	2	1
115	1	2	3
116	1	2	2
117	4	5	2
118	2	2	2
119	2	2	2
120	2	2	2
121	1	3	1
122	2	3	2

Explanation

1 Superior
2 Very Good
3 Good
4 Fair
5 Weak

TEACHERS' TRAITS

1. Adaptability.
2. Attractiveness, personal appearance.
3. Breadth of interest (interest in community, interest in profession, interest in pupils).
4. Carefulness (accuracy, definiteness, thoroughness).
5. Considerateness (appreciativeness, courtesy, kindliness, sympathy, tact, unselfishness).
6. Co-operation (helpfulness, loyalty).
7. Dependability (consistency).
8. Enthusiasm (alertness, animation, inspiration, spontaneity).
9. Fluency.
10. Forcefulness (courage, decisiveness, firmness, independence, purposefulness).
11. Good judgment (discretion, foresight, insight, intelligence).
12. Health.
13. Honesty.
14. Industry (patience, perseverance).
15. Leadership (initiative, self-confidence).
16. Magnetism (approachability, cheerfulness, optimism, pleasantness, sense of humor, sociability, pleasing voice, wittiness).
17. Neatness (cleanliness).
18. Openmindedness.
19. Originality (imaginativeness, resourcefulness).
20. Progressiveness (ambition).
21. Promptness (dispatch, punctuality).
22. Refinement (conventionality, good taste, modesty, morality simplicity).
23. Scholarship (intellectual curiosity).
24. Self-control (calmness, dignity, poise, reserve, sobriety).
25. Thrift.

Commonwealth Teacher-Training Study by Charters and Waples c 1929, page 18.

COLLEGE OF THE PACIFIC
STOCKTON, CALIFORNIA
OFFICE OF THE APPOINTMENT SECRETARY

A study of the teachers trained at the College of the Pacific is being made by Emma Fentzling, a graduate student in the field of education, under the sponsorship of Dr. J. William Harris, dean of the School of Education at the College of the Pacific.

The data called for on these blanks will be used in group tabulations which will not identify either institutions or individuals.

You will facilitate this study if you will answer the questions on the enclosed data sheet in regard to the teacher who has taught for you or who is at the present time teaching under your supervision.

Your cooperation will be greatly appreciated.

Cordially yours,

Clarence Scola
Placement Secretary

COLLEGE OF THE PACIFIC
STOCKTON, CALIFORNIA
Office of the Placement Secretary

.....

.....
has been engaged in educational work under your jurisdiction.

We shall appreciate your classifying this person, using whichever of the following terms you deem appropriate: Superior, very good, good, fair, weak.

BERNIECE FIOLA,
Placement Secretary

Personality	Energy and persistence.....
Knowledge of subjects.....	Culture and refinement.....
Judgment	Community interest and standing .....
Health and vigor	Ability in discipline .....
Personal appearance.....	Skill as an instructor.....
Power of oral expression	Influence on students .....

General report:

Length of my period of observation:.....

Signed.....

Title .....

Dated at .....

Note: Please add personal comments which you would prefer having omitted from data sent to persons making inquiries, but which will assist us in properly classifying registrant.

(This report will be treated as confidential)

Case Number_____

Name_____

Year of test_____

Year of graduation_____

Major subject_____

Type of Credential_____

Quintile from entrance test_____

G.P.A. from four year record_____

Teaching success from principal's report_____

	Major Professor		Cadet Supervisor	Principal	
1. Personality					
2. Scholarship, Knowledge					
3. Judgment					
4. Health, Vigor					
5. Personal appearance					
6. Oral expression					
7. Energy, persistence					
8. Culture, refinement					
9. Community interest					
10. Ability in discipline					
11. Skill as an instructor					
12. Influence on students					
Total rating					
Quintile					

1. Superior, 2. very good, 3. good, 4. fair, 5. weak.

BIBLIOGRAPHY

BIBLIOGRAPHY

Books

Charters, and Waples, Commonwealth Teacher-Training Study. Edited by University of Chicago Press, 1929.

Downey, June E., The Will Temperament and Its Testing. Edited by The World Book Co., 1922

Ellefson, E., Predicting Scholarship, Univ. of Calif., 1928.

Symonds, Percival M., Measurement in Secondary Education. Edited by The Macmillan Co., 1927.

Wood, Ben D., Measurement in Higher Education. Edited by The World Book Co., 1923.

Periodicals

Armentrout, W. D., "The Rating of Teachers by Training Teachers and Superintendents", Elementary School Journal, XXVIII, No. 7.

Bowman, Earl C., "The Problem of the Early Prognosis of Teaching Success", Educational Administration and Supervision, Feb., 1931.

Breckenridge, Elizabeth, "A Study of the Relation of Preparatory School Records and Intelligence Test Scores to Teaching Success", Educa. Admin. and Superv., Dec., 1931.

Burr, S. E., "Why the I. Q. Needs a New and More Descriptive Name", The Nations' Schools, Apr., 1931.

Clement, Evelyn, "An Evaluation of Teacher-Training", Educa. Admin. and Superv., Feb., 1932.

Conrad, H. S. "The Effect of the Personal Equation", Journal of Educational Psychology, Feb., 1932.

Dexter, Emily S., "Intelligence Test Score and Major Subject", School and Society, XXX, No. 780.

Freeman, Frank S., "Elusive Factors Tending to Reduce Correlations Between Intelligence Test Ranks and College Grades", School and Society, XXIX, No. 755.

Gillis, Frances M., "Correlates of Intelligence in College Students", School and Society, Aug. 22, 1931.

Gist, Arthur S., "Important Points of View in Teacher-Training", Educa. Admin. and Supery., Apr., 1931.

Hamrin, S. A., "A Comparative Study of Ratings of Teachers-in-Training and Teachers-in-Service", Elementary School Journal, XXVIII, No. 1.

Hertzberg, O. E., "New Rigid Entrance Requirements for Teacher-Training Institutions", School and Society, Nov. 21, 1931.

Hughes, W. H., "Refining the Estimates of Personal Qualities", The Nation's Schools, Feb., 1931.

Hughes, W. H., "Why Intelligence Test Scores are Not More Highly Predictive of School Success", Educa. Admin. and Supery., XII, No. 1.

Knight, F. B., "Qualities Related to Success in Teaching", Teachers College Contribution to Education, New York, No. 120.

McCory, John R., "A Study of the Relation Between Ability and Achievement", Educa. Admin. and Supery., XII, No. 7.

Middlebrook, L. Ruth, "A Modest Proposal", School and Society, Oct. 10, 1931.

Neel, Mary O., and Mead, A.R., "Correlations Between Certain Group Factors in Preparation of Secondary School Teachers", Educa. Admin. and Supery., Dec., 1931.

Newmark, David, "Students' Opinions of Their Best and Poorest Teachers", Elementary School Journal, XXIX, No. 8.

Peterson, Oda K., and Cook, Wm. A., "Score Cards and Rating Sheets in Teacher Training", Educational Method, IX, No. 6.

Ruediger, W. C., "Rating Teachers", Sch. and Soc., XX, No. 505.

Rugg, Harold O., "Is the Rating of Human Character Practical?", Journal of Educational Psychology, XII, No. 8 and No. 9, and XIII, No. 2.

Rugg, Harold O., "Self-improvement of Teachers Through Self-rating", Elemen. Sch. Jr., May, 1920.

Thayer, V. T., "Teacher Rating in the Secondary School", Educa. Admin. and Superv., XII, No. 6.

---"Concerning Marks", School and Society, Oct. 10, 1931.

---"Training Requirements for Newly Appointed Teachers", Research Bulletin of the National Education Association, Jan. 1932.