

THE UNIVERSITY OF MINNESOTA

GRADUATE SCHOOL

Report
of
Committee on Thesis

The undersigned, acting as a Committee of the Graduate School, have read the accompanying thesis submitted by **Manuel Lingad Carreon** for the degree of **Master of Arts.**

They approve it as a thesis meeting the requirements of the Graduate School of the University of Minnesota, and recommend that it be accepted in partial fulfillment of the requirements for the degree of **Master of Arts.**

W. G. Huggerty
.....
Chairman

H. L. Dealey
.....

F. Kuhlman
.....

May 27 *1928*

THE UNIVERSITY OF MINNESOTA

GRADUATE SCHOOL

Report
of
Committee on Examination

This is to certify that we the undersigned, as a committee of the Graduate School, have given Manuel Lingad Carreon final oral examination for the degree of Master of Arts . We recommend that the degree of Master of Arts be conferred upon the candidate.

Minneapolis, Minnesota

May 26, 1921

M. E. Hepperty
Chairman

J. B. Sears

Ross L. Juncy

Raymond W. Sieis

Philip K. Kasper
W. J. Van Wagenen

L. G. Koos
A. V. Storm

THE RÔLE OF STANDARDIZED TESTS

IN

PHILIPPINE PUBLIC SCHOOL ADMINISTRATION

A THESIS

SUBMITTED TO THE GRADUATE FACULTY

OF THE

UNIVERSITY OF MINNESOTA.

BY

MANUEL L. CARREON

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

FOR THE

DEGREE OF

MASTER OF ARTS

JUNE
1921

UNIVERSITY OF
MINNESOTA
LIBRARY

MOM
9C23
8

TABLE OF CONTENTS

PART I

STATEMENT OF THE PROBLEM AND THE TESTING MOVEMENT

CHAPTER I

Page

INTRODUCTION - STATEMENT OF THE PROBLEM

1

- (a) Purpose and General Scope of the Study
- (b) Sources of Material
- (c) Methods Followed in the Study

CHAPTER II

THE TESTING MOVEMENT

5

- (a) Brief History of the Movement
 - (1) Development of Achievement Tests
 - (2) Development of Intelligence Tests
 - (3) The Triumph of the Testing Movement
- (b) Types of Tests and Scales
- (c) Recent Tests and Scales
- (d) Tendencies in the Use of Measurements
- (e) Agencies of the Testing Movement
 - (1) Research and Efficiency Bureaus
 - (2) Educational Associations and Publications
- (f) Summary

290328

OCT 6 1922 Ed. 1.35.

TABLE OF CONTENTS

PART II

THE RELATION OF STANDARDIZED TESTS TO PUBLIC SCHOOL
ADMINISTRATION

CHAPTER III

	Page
STANDARDIZED TESTS IN BUSINESS AND SCHOOL ADMINISTRATION	24
(a) The Fundamental Principles of any Form of Administration	
(b) The Value of Standards in the Applica- tion of these Principles to Business Administration	
(c) The Parallel Use of Standard Measure- ments in Educational Administration	
(d) Summary	

CHAPTER IV

THE APPLICATION OF TESTS AND MEASUREMENTS TO AMERICAN
EDUCATIONAL PROBLEMS

	35
(a) Some Outstanding Problems in American School Administration	
(b) The Problem of Grading and Promotion	
(c) The Problem of Retardation and Elimination	
(d) Methods of Instruction and the Curriculum	
(e) Educational and Vocational Guidance	

TABLE OF CONTENTS

- (f) School Supervision and Organization
- (g) Summary

CHAPTER V

Page

THE APPLICABILITY OF STANDARDIZED TESTS TO PHILIPPINE

PUBLIC SCHOOL CONDITIONS

53

- (a) The Evolution and Administration of the
Philippine Public School System
- (b) Measurements in Solving Philippine Public
School Problems
- (c) Studies on the Applicability of Tests to
Filipinos
 - (1) Study on Filipino Students in Minnesota
 - (2) Study on Filipino Students in Ohio
 - (3) Study on Filipino Teachers in the
Philippines
- (d) Summary

CHAPTER VI

SUMMARY AND CONCLUSIONS

74

- (a) The Testing Movement in America and its
Educational Significance
- (b) Tests and Measurements in American Public
School Administration

TABLE OF CONTENTS

(c) The Rôle of Standardized Tests in Philippine
Public School Administration

LIST OF FIGURES

	Page
Figure 1	Efficiency of Supervision 32
Figure 2	Overlapping in the Mental Ages of Children in the First, Fifth, and Ninth Grades 38
Figure 3	Percentile Distribution of Scores for Men and Women - Provincial Group 66
Figure 4	Percentile Distribution - Baguio (City) Group 68
Figure 5	Profile Graph Showing Average Scores in Each Test 70

LIST OF TABLES

Table I	Scores and Frequencies, Minnesota Group 60
Table II	Scores and Frequencies, Ohio Group, Army Alpha 62
Table III	Scores and Frequencies, Ohio Group, Ohio State University Test 63
Table IV	Scores and Frequencies, Provincial Group 65
Table V	Average Scores by Provinces 66-67
Table VI	Average Scores by Educational Attainments 67
Table VII	Scores and Frequencies, Baguio Group 67-68
Table VIII	Average Scores by Educational Attainments 69

PART I

STATEMENT OF THE PROBLEM AND THE TESTING MOVEMENT

CHAPTER I

INTRODUCTION--STATEMENT OF THE PROBLEM

(a) Purpose and General Scope of the Study

The study aims to show: (1) the relation of standardized tests and measurements to the problems of public school administration; and (2) the application of this relationship to Philippine educational theory and practice. Tests and measurements, as used here in their relation to public school administration, mean the achievement and the intelligence tests. No effort has been made to include in this study other tests and measurements because these have not yet been well standardized and widely used. It is not the purpose of this dissertation to enter into a lengthy discussion of statistical terminology. Throughout the entire discussion, the terms "test", "measurement", and "scale" will be used. Measurement and test refer to any device for measuring accomplishment and native ability. If they measure the former, they are called achievement tests; if the latter, they are known as intelligence or mental tests. Here and there in the study the word "scale" is also used. A scale is a test or measurement in which a common unit is used. It differs from the ordinary test in that the latter simply measures the number of tasks or problems done without regard to a common unit of measurement. Thus a scale is merely a development of a test, be it achievement or endowment. The study, therefore, treats of

standardized measurements of both performance and intelligence, either as tests or as scales, as the case may be.

(b) Sources of Material

Literature on tests and measurements published in educational periodicals and magazines is the main source of material used for the first part of the dissertation. Articles in the Twelfth, Fifteenth, and Seventeenth Yearbooks of the National Society for the Study of Education, the Journal of Educational Research, Elementary School Journal, Educational Review, and Teachers College Record have given the writer most of the current material on the subject. A few extracts have also been taken from Educational Administration and Supervision, the proceedings of the National Education Association, and a few other school publications. Then the courses in Mental and Educational Diagnosis and Educational Administration taken both at the University of Nebraska and at the University of Minnesota have helped to throw light on the problems taken up here. The discussions in the seminar and the conferences with his adviser and instructors in the latter institution have been no less helpful to the writer in the organization of his material.

For the second part of the study, three sources have been available: (1) results of Delta 2 tests administered to Filipino students of Minneapolis and St. Paul, Minnesota, in the winter quarter of 1920-1921; (2) results of Army Alpha and Ohio State University Group tests furnished by the department of psychology at Ohio State University, Columbus, Ohio from the data obtained in classes where there are Filipino students; and (3) results of

the Otis Advanced Intelligence Scale given to Filipino teachers during the 1919 summer convention of teachers and superintendents at Baguio, Philippines, and to teachers at the normal institutes in four provinces (Batangas, Laguna, Rizal, and Tayabas) of the same year.¹

(c) Methods Followed in the Study

The only method that might be feasibly and profitably used for the first part of the study is the documentary method in the form of research or reference. Along with the courses in Mental and Educational Diagnosis and Educational Administration, and the reports and discussions in the seminar in educational psychology, the writer has taken hold of various tests and scales, studied them rather carefully, and read articles on results obtained from them and on the relation of these results to public school administration with a view to determining the relative values of these measurements when it comes to applying them to the Philippines. In making use of the literature on measurements in school administration, every effort has been made to discriminate and sort out such facts as have a direct bearing on Philippine school conditions. The aim in this selective process has been the treatment of only such school problems as both the American and the Philippine administrator would meet.

The second part of the study has necessitated both practical field work and the application of statistical methods.

1. This part of the study was made possible largely through the co-operation of the Filipino students of Minnesota, the department of psychology at Ohio State University, Columbus, Ohio, and the Second Assistant Director of Education, Manila.

Tests were given to twenty-two Filipino students residing in Minneapolis and St. Paul, scores recorded, and results compared and tabulated. The entire work of testing, scoring and interpreting results was done by the writer. The department of psychology at Ohio State University, Columbus, Ohio has been kind enough to furnish the material on the Army Alpha and Ohio State University intelligence tests. Twelve Filipino students were tested on the Army Alpha and nineteen on the Ohio State University test. The material on the Otis test has been secured through the Assistant Director of Education, Manila, Philippines.² In the last two investigations, only the organization of material, tabulation and a few minor computations, and the interpretation of results were left to the writer. A full discussion of these studies is found in Chapter V of the dissertation.

2. The Otis tests were given by Abram van Heyningen Hartendorp of the U. S. Army in the Philippines.

CHAPTER II

THE TESTING MOVEMENT

(a) Brief History of the Movement

Statistical methods to measure educational achievements and mental abilities are only a recent development in the field of educational science. According to Ayres, measurements in education are fifty years old if traced to the oldest beginnings recorded, twenty-five if reckoned from Dr. Rice's first experiment, and ten if considered only from the earliest efforts of Thorndike to construct his handwriting scale.¹ At first methods were borrowed from economics, anthropology, and the physical sciences. Later purely educational measurements began to evolve. And by the time Dr. Thorndike first published his handwriting scale, a definite technique had been developed.

(1) Development of Achievement Tests

As early as 1864, Rev. George Fisher, of the Greenwich Hospital School, saw the need and possibility of quantitative measurement in education. By means of a "Scale-Book", in which different numerical values ranging from 1 (the highest) to 5 (the lowest) are assigned to different degrees of proficiency in the various subjects - writing, spelling, mathematics, navigation, Scripture, knowledge, grammar and composition, French, general history, drawing, and practical science, he was able to measure achievements in these subjects.² Unfortunately, nobody

1. Ayres, L. P. History and Present Status of Educational Measurements. Seventeenth Yearbook of the National Society for the Study of Education. Part II, 1918, p. 9.

2. Ibid., pp. 9-10 (Also reported by E. B. Chadwick in the

seemed to have paid much attention to this pioneer work, and consequently, Rev. Fisher's efforts did not produce lasting results.

The development of statistical methods for the quantitative study of what seemed to be entirely qualitative material, and the work of Sir Francis Galton in experimental and educational psychology gave a new impetus to the testing movement.³ It was at this juncture that Dr. Rice, the so-called inventor of educational measurement, appeared on the scene. Educated in Germany under the influence of eminent psychologists at Jena and Leipsic, he returned to America in 1894 with the new idea. This new idea was the setting up of standards for achievement in the various subjects. He started out with spelling by making up a list of 50 words, tested pupils from school to school, and made the startling report that children taking spelling thirty minutes a day were no better off than those taking it only half that long. The present-day conflict between the investigator and the formalist in education began. Gradually, Dr. Rice won support from the more thoughtful educators, one of whom was Professor Hanus of Harvard.⁴

Following the footsteps of Rice, Professor Thorndike, about ten years later, began to experiment with tests and scales. His handwriting scale was first published in 1910.⁵ The publication of this scale marks the real beginning of scientific measurements

Museum, a Quarterly Magazine of Education, Literature and Science, Vol. III, 1864).

3. Ibid., pp. 10-11.

4. Ibid., p. 12.

5. Thorndike, E. L. Handwriting, Teachers College Record, Vol. II, March, 1910, pp. 1-93.

in the field of education. Two years later, the Hillegas composition scale appeared,⁶ and since then tests and scales in different subjects of the curriculum have been devised and standardized. A full treatment of those most highly standardized and widely used is given later in this chapter in the list of achievement tests.

(2) Development of Intelligence Tests

The history of mental tests goes back to only about a decade ago. It was through the untiring efforts of the French psychologist, Alfred Binet, that intelligence tests first came into use. At first connected with institutions for mental defectives, Binet conceived the idea of devising certain tests for the mental faculties. Through the co-operation of his colleague, Dr. Simon, he first proposed his tests in 1905, replaced them with a more systematic formulation in 1908, and effected final revisions in 1911. Binet's tests aim to measure general intelligence.⁷ As left by him, there are fifty-four tests in all, ranging from year 3 to adult.⁸

The Binet-Simon tests (as they are generally known) mark the beginning of mental testing. They have met such a success and won such a popularity in different countries that several formulations of the original scale have appeared from time to

6. Hillegas, M. B. A Scale for the Measurement of Quality in English Composition by Young People. Teachers College Record, Vol. XIII, Sept. 1912, pp. 331-384.

7. Terman, L. M. "The Measurement of Intelligence", pp. 41-43; 1916, Houghton Mifflin Company.

8. Binet, A. Nouvelles Recherches sur la Mesure du Niveau Intellectuel chez les Enfants d'École. L'Année Psychologique (1911), Vol. 17, pp. 145-201.

time. The most important of these revisions and extensions in America are those by Goddard,⁹ Terman,¹⁰ and Kuhlmann.¹¹ Although differing from the Binet-Simon intelligence scale in that scores are recorded in terms of points continuously from 1 to 20 instead of by chronological and mental ages, and that different norms apply to different racial, social, environmental, and linguistic groups, the Yerkes-Bridges Point-Scale is also patterned after the Binet-Simon scale.¹² Not only did the Binet-Simon intelligence scale lead to its various formulations and modifications, but it also paved the way for other attempts in mental testing.

With the breaking out of the World War and America's entering in the great conflict, a new movement in mental testing began. Up to that time, practically nothing had been done in the field of group testing. But when America entered the Great War, a need for the proper classification of army men to utilize "brain power" as much as possible was immediately felt. The psychologists saw this opportunity, and in the spring of 1917 they offered their services to the government.¹³ Two varieties of group tests were devised, the alpha for literates and the beta

9. Goddard, H. H. The Binet-Simon Measuring Scale for Intelligence, Revised. Training School Bulletin (1911), Vol. III, pp. 56-62.

10. Terman, L. M. "The Measurement of Intelligence", Stanford Revision and Extension of the Binet-Simon Intelligence Scale. 1916, 348 pages. Houghton Mifflin Company.

11. Kuhlmann, F. A Revision of the Binet-Simon System for Measuring the Intelligence of Children. Monograph Supplement of Journal of Psycho-Asthenics (September, 1912), 41 pages.

12. Yerkes, Bridges, and Hardwick. "A Point Scale for Measuring Mental Ability"; 168 pages, 1915, Warwick and York, Inc.

13. Yerkes and Yoakum. "Army Mental Tests". Introduction, p. viii. 1920. Henry Holt and Company.

for illiterates. Only special cases were given individual examinations.¹⁴ The success of the army tests gave the greatest impetus to the mental test movement. Several adaptations of these tests followed as in the case of the Binet-Simon tests. One finds such adaptations in Haggerty's Delta 1 and 2,¹⁵ Terman's Group Test of Mental Ability,¹⁶ the National Intelligence Tests,¹⁷ Otis' Group Intelligence Examination,¹⁸ Thurstone's Psychological Examination,¹⁹ Pressey's Cross-Out Tests,²⁰ Smith's Graded Intelligence Tests,²¹ and numerous others.

(3) The Triumph of the Testing Movement

As in all new movements, progress in educational measurement was being obstructed by certain adherents of the old school-superintendents, supervisors, principals, and teachers who could not see the wonderful possibilities of the testing movement. They claimed there was no use of measuring things unknown and unmeasurable. They contended mental constituents are too intangible to measure. They branded the new movement as one dealing merely with the formal and mechanical aspects of education.

14. Ibid., pp. 15-19.

15. Haggerty, M. E. Intelligence Examination, Delta 1 and Delta 2. 1920. World Book Company.

16. Terman, L. M. Group Test of Mental Ability. 1920. World Book Company.

17. Haggerty, Terman, Thorndike, Whipple, and Yerkes. National Intelligence Tests. 1920. World Book Company.

18. Otis, A. S. An Absolute Point Scale for the Measurement of Intelligence. Journal of Educational Psychology, May and June, 1918, reprints.

19. Thurstone, L. L. Psychological Examination. 1919. Carnegie Institute of Technology, Form 1481.

20. Pressey, S. L. A Brief Group Scale of Intelligence for Use in School Surveys. Journal of Educational Psychology, Vol. XI, No. 2, Feb. 1920, pp. 89-100.

21. Smith, F. O. Graded Intelligence Tests. Revised from University of Montana Bulletins Nos. 233, 234, 235, and 236, 1920.

Raising the very same objection, Superintendent Horn writes in part:

"It should furthermore be kept in mind that there are many things about a school system which can never be definitely measured or stated with mathematical accuracy. Just where the line is to be drawn between the measurable and the non-measurable elements that enter into a school is a matter concerning which there is much difference of opinion. In other words, the element of opinion enters to some extent even into the matter of the possibility of measurement.

"For instance, it is an undoubted fact that any man can go into a city and count the school houses or the number of the desks. Any man can find out the number of teachers employed. Any man can count for himself the number of pupils present in a given room.

"It takes no particular ability to enable an inquirer to find out just how much money is being spent. If the schools spend nine hundred thousand dollars in one year and a million dollars the next year, any one can deduce the fact that they spent one hundred thousand dollars more the second year than the first year.

"On the other hand, after a comparatively few such facts have been definitely ascertained, we come to subjects that cannot be measured in mathematical terms, and concerning which there are no definite standards. In this realm ideals are not always definitely established and opinions are almost certain to vary widely."

Then he goes on to point out the impossibility of measuring the most essential elements making for an efficient school system as follows:

"A school that turns out manly, honorable, self-reliant boys and womanly, efficient girls is likely to be at least a fairly good school, no matter what it may do otherwise. A school that fails to turn out such pupils can hardly be considered a good one, no matter what it may do for its pupils in the way of reading, or writing, or arithmetic. And yet these very things, which may decide between the success or failure of the school, are matters which it is almost impossible to estimate accurately, and concerning which there may be a wide amount of honest difference of opinion."²²

But the success of the Army Tests²³ and the part played by group tests in the several school surveys made throughout the United

22. Horn, P. W. Report of Supplementary Survey of Portland Public Schools, pp. 6-7. April, 1917.

23. Quotations from statements of officers (Yerkes and

States drove away gradually the cloud of pessimism and assured a prominent place for tests in the realm of public school administration.

(b) Types of Tests and Scales

Tests and scales now in existence may be roughly divided as follows:

(1) Achievement Tests - Tests on subject-matter and information learned in school.

(2) Intelligence Tests - Individual or group, special or general.

(3) Tests of Aesthetic Functions.

(4) Tests of Ethical, Volitional, and Emotional Qualities.

The achievement tests are the most numerous of all. Several are still in preparation, some will be off the press in a short time, and others have already appeared recently. The following are some of the most highly standardized and widely used:²⁴

A. Elementary School Tests

Yoakum, "Army Mental Tests", pp. 13-15):

"Officers and men should be given a psychological examination as a matter of routine.

"The results of the psychological examinations are fully borne out by actual observation of the abilities and the capacity of various officers in the performance of duties assigned to them. I do not mean by this that these tests are an absolute gauge, but I do mean that they are an absolute guide, and that given the practical tests we are enabled to arrive at the best possible determination of ability to meet the requirements of the service."

"All enlisted men sent to Officers' Training Schools from this camp are inspected as to their military appearance and bearing, and their knowledge of the elementary duties of a soldier; they are given a physical examination, a mental examination, and the psychological examination. If they do not rate A or B in this examination they are rejected."

24. Selected with the assistance of Dr. M. J. Van Wagenen,

I. Arithmetic - Fundamental Operations and Reasoning

- a. Cleveland Survey Tests²⁵
- b. Curtis Standard Research Tests, Series B²⁶
- c. Monroe's Diagnostic Tests²⁷
- d. Woody's Arithmetic Scales²⁸
- e. Buckingham's Reasoning Tests²⁹

II. Geography

- a. Hahn-Lackey Geography Scale³⁰

III. Handwriting

- a. Ayres' Scale for Measuring the Handwriting of School Children³¹
- b. Ayres' "Gettysburg Edition"³²

University of Minnesota, from lists given in: (1) Seventeenth Yearbook of the National Society for the Study of Education, Part II, pp. 71-104; (2) Strayer and Engelhardt, "The Classroom Teacher", pp. 205-211, 1920, World Book Company; and (3) Journal of N. E. A. for April, 1921, Vol. X, pp. 79-80.

25. Judd, C. H. Measuring the Work of the Public Schools. The Survey Committee of the Cleveland Foundation, Cleveland, Ohio, 1916, 290 pages.

26. Curtis, S. A. Third, Fourth, and Fifth Annual Accounting of the Curtis Standard Research Tests, 1913-1916; also Capacity, Ability, and Performance in Relation to Standard Scores and Summary of Tabulations. Bulletin No. 4, Department of Co-operative Research, Detroit, Mich., 111 pages.

27. Monroe, W. S. Existing Tests and Standards. Seventeenth Yearbook of the National Society for the Study of Education, Part II, p. 75.

28. Woody, C. Measurements of Some Achievements in Arithmetic. School and Society, Vol. IV, Aug. 19, 1916, pp. 229-303.

29. Buckingham, B. R. Notes on the Derivation of Scales in School Subjects, with Special Application to Arithmetic. Fifteenth Yearbook, National Society for the Study of Education, Part I, 1916, pp. 23-40.

30. Hahn and Lackey. Geography Scale. Wayne State Normal School, Wayne, Nebraska.

31. Ayres, L. P. A Scale for Measuring the Quality of Handwriting of School Children. Russel Sage Foundation, Bulletin E 113, 1912, 16 pages.

32. Ayres, L. P. A Measuring Scale for Handwriting. "Gettysburg Edition". Russel Sage Foundation, 1917. New York City.

c. Freeman's Handwriting Scale³³

d. Thorndike's Scale³⁴

IV. History

a. Van Wagenen's U. S. History Scales³⁵

V. Language and Composition

a. Nassau County Supplement to the Hillegas Scale³⁶

b. Thorndike's Extension of the Hillegas Scale³⁷

c. The Trabue Completion-Test Language Scales³⁸

d. Hudelson's English Composition Scale³⁹

VI. Reading - Silent and Oral

a. Gray's Silent Reading Tests⁴⁰

b. Thorndike's Visual Vocabulary Scales⁴¹

c. Thorndike's Scale Alpha and Alpha 2 for Measuring the Understanding of Sentences⁴²

33. Freeman, F. N. An Analytical Scale for Judging Handwriting. Elementary School Journal, Vol. XV, April, 1915, pp. 432-441.

34. Thorndike, E. L. Handwriting. Teachers College Record, Vol. II, March, 1910, pp. 1-93.

35. Van Wagenen, M. J. Historical Information and Judgment in Pupils of Elementary Schools. Teachers College, Columbia University Contributions to Education, No. 101, 1919.

36. Trabue, M. R. Supplementing the Hillegas Scale. Teachers College Record, Vol. XVIII, Jan., 1917, pp. 51-84.

37. Thorndike, E. L. Preliminary Extension to the Hillegas Scale. Bureau of Publications, Teachers College, Columbia University, New York City.

38. Trabue, M. R. Completion-Test Language Scales. 119 pages. Bureau of Publications, Teachers College, Columbia University, New York City.

39. Hudelson, E. English Composition Scale. 46 pages. 1921. World Book Company.

40. Gray, W. S. Methods of Treating Reading. Elementary School Journal, Vol. XVI, Jan., 1916, pp. 231-246 and Feb., 1916, pp. 281-298.

41. Thorndike, E. L. Visual Vocabulary Reading Test. Bureau of Publications, Teachers College, Columbia University, New York City.

42. Thorndike, E. L. An Improved Scale for Measuring Ability

d. Haggerty's Sigma 1 and 3⁴³

e. Gray's Oral Reading Test⁴⁴

VII. Spelling

a. Ayres' Spelling Scale⁴⁵

b. The Iowa Dictation Exercise and Spelling Tests⁴⁶

c. Buckingham's Spelling Scale⁴⁷

B. High School Tests

I. Mathematics

a. Hotz's First-Year Algebra Scales⁴⁸

b. Rugg and Clark Standardized Tests in First-Year Algebra⁴⁹

c. Minnick's Geometry Tests⁵⁰

II. Foreign Languages

a. Henmon's Latin Tests⁵¹

III. History

a. Sackett's Scale in Ancient History⁵²

in Reading. Teachers College Record, Vol. XVI, Nov., 1915; and Vol. XVII, Jan., 1916, pp. 40-67.

43. Haggerty, M. E. Reading Examination. 1920 and 1921. World Book Company.

44. Vide note 40, page 13.

45. Ayres, L. P. A Measuring Scale for Ability in Spelling. Bulletin E 139, 56 pages, Russel Sage Foundation, New York City.

46. Asbaugh, E. J. Extension Division, University of Iowa.

47. Buckingham, B. R. Spelling Ability: Its Measurement and Distribution. Teachers College Contributions to Education, No. 59, 1913, 116 pages.

48. Address Bureau of Publications, Teachers College, Columbia University.

49. Rugg, H. O. and Clark, J. R. Standardized Tests and the Improvement of Teaching in First-Year Algebra. School Review, Vol. XXV, Feb. and Mar., 1917, pp. 113-132 and 196-213.

50. Address J. H. Minnick, University of Pennsylvania, Philadelphia, Pennsylvania.

51. Henmon, V. A. C. The Measurement of Ability in Latin. Part I, Vocabulary. Journal of Educational Psychology, Vol. VIII, Nov. 1917, pp. 515-538.

52. Sackett, L. W. A Scale in Ancient History. Journal of

IV. Science

a. Starch's Tests in Physics⁵³b. Powers' Tests in Chemistry⁵⁴

Intelligence tests are individual or group and special or general. When they are classified according to the manner in which they are given, they are either individual or group. The following are the most satisfactory individual and group tests:

I. Individual⁵⁵

a. Stanford-Binet Intelligence Scale

b. Goddard Revision of the Binet Scale

c. Yerkes-Bridges Point Scale

II. Group⁵⁶

a. Army Alpha and Beta

b. Terman's Group Test of Mental Ability

c. Otis Group Intelligence Scale

d. Haggerty Delta 1 and 2

e. National Intelligence Tests

f. Pressey Primer and Cross-Out Scales

g. Thurstone's Psychological Examination

h. Smith's Graded Intelligence Tests

i. Dearborn Group Tests of Intelligence and General Examinations⁵⁷

Educational Psychology, Vol. VIII, May, 1917, pp. 284-293.

53. Starch, D. Test for Physics. University of Wisconsin, Madison, Wisconsin.

54. Only recently devised.

55. Vide p. 8.

56. Vide p. 9.

57. Dearborn, W. F. Group Tests of Intelligence, Series 1. General Examinations 1, 2, and 3, for Grades I-III. 1920. J. B. Lippincott, Philadelphia.

j. Miller High School Test⁵⁸

When classified according to what they test, intelligence examinations are either special or general. Of special characteristic tests, we have Stone's reasoning tests,⁵⁹ J. B. Herring's tests for scientific thinking,⁶⁰ Van Wagenen's United States History tests for information, thinking, and character judgment,⁶¹ and a similar attempt made by E. U. Rugg in general history.⁶² All the tests listed in the first classification are general intelligence tests, whether individual or group.

There are not many tests devised so far to measure aesthetic functions. The problem of producing such tests has been no easy matter. A pioneer in this field is Thorndike's scale for drawing. Devised in 1913, it consists of a series of drawings arranged in order of merit as determined by competent judges.⁶³ H. O. Rugg's Scale for Measuring Freehand Lettering for Use in Secondary Schools and Colleges is another example. This consists of a series of eight samples of freehand lettering, arranged in order of increasing merit. The scale may be used in measuring the efficiency of a student's work in freehand lettering.⁶⁴ Pintner is measuring the same functions in his tests of children's

58. Miller, W. S. Tests for High School Pupils. Form A. Standardized by Administrative Section of High School Conference, University of Minnesota.

59. Stone, C. W. Standardized Reasoning Tests in Arithmetic and How to Utilize Them. Bureau of Publications, Teachers College, Columbia University. 1916. 24 pages.

60. Journal of Educational Psychology, Dec., 1918, p. 535.

61. Vide p. 13.

62. School Review, Vol. XXVII, 757, Dec., 1919.

63. The Seventeenth Yearbook, National Society for the Study of Education, 1918, Part II, p. 78.

64. Ibid., p. 100.

appreciation of pictures.⁶⁵ C. C. Peters, of Ohio Wesleyan, is now working on a scale to measure aesthetic appreciation in literature.⁶⁶ Along the same line is the recent article by Abbott and Trabue in Teachers College Record for March, 1921, on measuring ability to judge poetry.⁶⁷ In music, Seashore's remarkable work has made it possible to make evaluations of some of the elements in music. The scale is called a musical talent chart, based upon the analysis of musical ability, and offers a graphic means of representing the pupil's musical ability.⁶⁸

The measurement of ethical, volitional, and emotional qualities is another difficult task. Two attempts have already been made along the line of ethical measurement. Upton and Chassell have devised a scale for measuring habits of good citizenship.⁶⁹ Very much like this is H. O. Rugg's score-card plan for rating and self-rating.⁷⁰ The same tendency is quite noticeable in teacher rating, of which Landsittel's score-card is an example.⁷¹ Volitional qualities have also been measured recently. The attempt of Miss Downey of the University of Wyoming to measure the will is a pioneer in the field.⁷² Even the most intangible elements of human nature have become the objective of tests and meas-

65. Pedagogical Seminary, Vol. XXV, June, 1918.

66. Educational Review, Vol. LXI, No. 2, Feb., 1921, p. 119.

67. Teachers College Record, Vol. XXII, No. 2, March, 1921, pp. 101-126.

68. The Seventeenth Yearbook, National Society for the Study of Education, Part II, 1918, p. 89. Also Educational Review, Vol. LXI, No. 2, p. 119.

69. Teachers College Record, January, 1919.

70. School Review, May, 1920.

71. School and Society, Vol. VI, 774, Dec. 29, 1917.

72. Downey, June E. The Will-Profile, a Tentative Scale for Measurement of the Volitional Pattern. University of Wyoming Bulletin No. 3, Vol. XVI, No. 4b, Nov., 1919.

urements. Attempts have recently been made to test the emotions.⁷³ Whether or not these attempts will meet with any success depends to a large extent on a consideration of outside or extraneous factors in measuring such intangible human elements and on the validity and reliability of the tests.

(c) Recent Tests and Scales

Recently new tests and scales have appeared. Most of these are achievement tests although some test intelligence, too. McCall's reading test,⁷⁴ Buckingham's extension of the Ayres' spelling scale,⁷⁵ Ruch's test for information in general science,⁷⁶ Downey's general science scale,⁷⁷ J. C. Chapman's tests in physics covering mechanics, sound, light, magnetism, and electricity,⁷⁸ Handschin's test for the purpose of diagnosing the type of linguistic capacity one possesses,⁷⁹ L. A. Wilkins' "predetermination" tests for French and Spanish,⁸⁰ S. B. Davis' (Pittsburgh) tests in history of the American colonial period,⁸¹ and Sackett's new scale in United States history⁸² are some new additions to the field of educational measurements.

Herring's tests of scientific thinking already referred to combine a science test with an intelligence test. Van Wagenen's

73. Literary Digest, Vol. 68, No. 23, Feb. 12, 1921.

74. Educational Review, Vol. LXI, No. 2, Feb., 1921, p. 122.

75. Ibid.

76. General Science Quarterly, 4, 257.

77. School Science and Math., 19, 228, March, 1919.

78. School Review, 27, Jan. and Dec., 1919.

79. Modern Language Journal, 3, 1-4, Oct., 1918.

80. Bulletin, High Points, N. Y., City Schools, Feb., Oct., 1919.

81. Educational Review, Vol. LXI, No. 2, Feb., 1921, p. 123.

82. Ibid.

elementary history tests measure not only information but also reasoning or thinking and character judgment. The fact that certain achievement tests are intelligence tests at the same time is indicated by Buckingham's attempt to show that tests of historical information are valuable indices of thought activity.⁸³

New tests have also been devised in vocational subjects: F. M. Leavitt's tests of manual arts;⁸⁴ Wardner's test for knowledge of tools;⁸⁵ Bowman and Trilling's tests in textiles and clothing;⁸⁶ and Murdoch's scale for certain elements of hand sewing.⁸⁷

(d) Tendencies in the Use of Measurements

Tests and measurements have been used in various ways. (1) Originally, standardized intelligence tests were given to mental defectives and dependents. (2) Later, these were applied to public school children. Achievement tests were also included. (3) Success in the public schools led to a more extensive use of tests in school surveys. The Denver, St. Paul, Nassau County (N. Y.), Cleveland, North Carolina and Virginia surveys all used tests and measurements. (4) Then tests have been used to diagnose school defects and offer remedies for improvement. (5) Tests are also being used to replace the college entrance examinations. The experiment at Columbia University is worth while mentioning here. The work of Thurstone at the Carnegie Institute of Technology is

83. Journal of Educational Research, March, 1920.

84. Leavitt, F. M. Standardized Measurement Scales in the Field of the Industrial Arts. Industrial Arts Magazine, Vol. VIII, pp. 132-138, April, 1919.

85. Educational Review, Vol. LXI, No. 2, Feb., 1921, p. 123.

86. Journal of Home Economics, Vol. XII, Nov., 1920, pp. 486-491.

87. Educational Review, Vol. LXI, No. 2, Feb., 1921, p. 123.

another example. (6) Finally, tests have been taken up by the business world, used for vocational placement and guidance.

More specialized uses have also been made of tests. (1) Whipple and a few others used tests and scales to select exceptionally capable children and find methods to treat such cases.⁸⁸ (2) Buckner's recent study shows a tendency to use tests to diagnose abilities and training of children as individuals, rather than in groups.⁸⁹ (3) Henman's recent study indicates a growing use of tests to measure the pupils' progress throughout the school year.⁹⁰ The use of tests in educational guidance is reported by Proctor.⁹¹ (5) Finally, Brooks has illustrated the use of tests in rural school administration.⁹²

(e) Agencies of the Testing Movement

(1) Research and Efficiency Bureaus

There are three classes of research and efficiency bureaus. (a) Those that are conducted by higher educational institutions (normal school or state university) for experimental and research work have very little connection with the schools. Good examples are found at the Kansas State Normal School, Emporia, Kansas, University of Illinois, and Indiana University. (b) There are also bureaus run by local school systems for their own benefits. Those of Los Angeles, Cleveland, Detroit, and New York City may be cited. (c) Then the state departments of public instruction have also started establishing such bureaus. Wisconsin gives

88. Ibid., p. 121.

89. Ibid., p. 122.

90. Journal of Educational Research, 1, 81.

91. Ibid., 1, 369.

92. Ibid., 1, 392.

the best illustration of this type.

(2) Educational Associations and Publications

Two other important agencies have been promoting the use of tests and measurements. One is the educational association. Devoted primarily to this purpose is the newly organized National Association of Directors of Educational Research. The National Academy of Science has given a great portion of its educational section to the program of measurements. The Annual Schoolmen's Week at the University of Pennsylvania and the Indiana University annual conference have both taken up the study of the same subject. The Department of Superintendence of the National Education Association has also joined in the movement. The National Society for the Study of Education, besides discussing the subject at two of its annual programs, has published two of its yearbooks on tests and measurements.

Another agency for the promotion of tests and measurements includes all publications on the subject. Such books as Thorndike's "Theory of Mental and Social Measurements", Rugg's "Statistical Methods Applied to Education", Starch's "Educational Measurements", Monroe, De Voss, and Kelly's "Educational Tests and Measurements", Monroe's "Measuring the Results of Teaching", Chapman and Rush's "Scientific Measurement of Classroom Products", Yerkes and Yoakum's "Army Mental Tests", Terman's "Intelligence of School Children" and "Measurement of Intelligence" - all have helped to popularize the testing movement. Several periodicals have also contributed to the progress of tests and measurements. Among the most prominent in this endeavor are the Journal of Educational Research (organ of the National Association of Direc-

tors of Educational Research), Journal of Educational Psychology, Journal of Applied Psychology, School and Society, School Review, Elementary School Journal, Teachers College Record, Educational Review, and Pedagogical Seminary.

(f) Summary

The brief history of tests and measurements clearly shows what a bright future lies ahead. Beset with pessimism and conservatism, the early pioneers labored hard. And they have won. At first it was considered futile to measure the intangible elements of human nature. But the testing movement was bound to prosper. Largely through the efforts of Dr. Rice and Professor Thorndike, the first step, the measurement of school achievements, was attained. Scales have been so highly perfected that today one measures handwriting with as much accuracy as the length of a board. While achievement tests were being developed, a new movement started, in the field of mental testing, through the ingenuity of the French psychologist, Binet. This movement marks the second step in the progress of testing. Although met with more and stronger opposition than the first step, mental testing made more progress than was at first expected. Face to face with school men who were not ready to recognize the possibility of measuring such an unknown thing as intelligence, the advocates of mental tests were not at all discouraged. They, too, have finally won. Now comes the third step - the measurement of the volitional and emotional elements of human nature. Miss Downey's pioneer work along volitional measurement and several attempts to measure emotions mark the beginning of a new era in tests and

measurements.

PART II

THE RELATION OF STANDARDIZED TESTS TO PUBLIC SCHOOL ADMINISTRATION

CHAPTER III

STANDARDIZED TESTS IN BUSINESS AND SCHOOL ADMINISTRATION¹

(a) The Fundamental Principles of any Form of Administration

In any kind of organization - commercial, industrial, philanthropic, governmental, or educational - there are certain common recognized principles of management, direction, and supervision. The directive and supervisory staff must perform the following fundamental tasks: (1) They must define the ends towards which the organization is striving; (2) they must co-ordinate the labors of all under them to attain those desired ends; (3) they must find the best methods of work and make the workers use these methods; (4) they must determine the qualifications necessary for the workers and see that only those that possess these qualifications are employed; (5) they must supply the workers with detailed instructions as to the work to be done, the standards to be reached, the methods to be employed, and the materials and appliances to be used; (6) they must furnish the necessary materials and appliances; and (7) they must place all the incentives to

1. The entire chapter is based largely on: (a) Bobbitt, F. The Supervision of City Schools, Twelfth Yearbook of the National Society for the Study of Education, Part I, 96 pages, University of Chicago Press, 1913; and (b) Spaulding, F. E. Application of the Principles of Scientific Management, Proceedings of National Education Association, 1913, pp. 259-279.

stimulate desirable effort on the part of the workers.

(b) The Value of Standards in the Application of these Principles to Business Administration

In recent years, scientific management in industry has been the subject of great interest in the field of economics and business administration. It has sometimes been called "taylorism", from a certain Mr. Taylor, an efficiency engineer, who studied the idea in the bicycle industry.²

When applied to business administration, the fundamental principles of any organization enunciated above reduce themselves to the following principles of scientific management:

(1) There must be definite qualitative and quantitative standards for the final output.

(2) There must be definite qualitative and quantitative standards for each stage of the process from the raw material to the ultimate product.

(3) The management must find the most efficient methods of procedure for actual service under actual conditions and make the workers use them.

(4) They must determine standard qualifications for the workers.

(5) They must keep the workers supplied with detailed instructions as to the work to be done, the standards to be reached, the methods and appliances to be used.

(6) They must supply the tools, materials, and appliances

2. Hoxie, R. F. Scientific Management and Labor Welfare. Trade Unionism in the United States, Chapter XII, pp. 296-325, 1917, D. Appleton and Company.

most effective for the work in hand.

(7) They must place all incentives to stimulate the workers to produce the largest output possible.

The first principle of scientific management, then, is the use of standards. Once standards are determined upon, it is easy to carry out the other principles. It is evident that no efficient method can be evolved unless by the use of definite standards the results of different methods can be measured. No "trial-and-error" or "hit-or-miss" method of discovering the best devices for efficiency and economy in management can be relied upon. The only scientific way of determining the quality and quantity of the product is to use standardized measurements. The qualifications of workers can be determined only by setting up certain standards to which the applicants must measure up to be admitted. While in service, the measurement of the results of their labors will determine to a large degree their efficiency and capacity for improvement. Such measurement will help solve the problem of promotion, transfer, or separation from the service. Again, no detailed instructions on the work to be done, the standards to be reached, the methods and appliances to be used can be given by the management without any definite knowledge of the actual conditions in the plant. And this can be done only by measuring the output produced. The value of standard measurements is also manifest when it comes to supplies (materials and appliances). Only definite standards will tell what quality and quantity of supplies are desirable. The same is true in the case of giving incentives to the workers. Only by an actual knowledge of their a-

chievement or work can the management determine who need the most help and in what particular line, and who need simply encouragement and other incentives to stimulate desirable effort.

Prof. Bobbitt, in his article entitled "Scientific Management Applied to City Schools", published in the Twelfth Yearbook of the National Society for the Study of Education, cites the following example in the world of material production on the value of definite plans and specifications as to the nature of the product to be turned out:³ In the rail industry, the size, shape, physical and chemical qualities, and tempering of the rails are first determined upon before any work is begun. The value of such definite instructions to the superintendent of the plant is to enable him:

(a) To organize all of the forces at his command, direct them, and supervise them in such a fashion as to secure just the product desired.

(b) To select the most suitable machinery for the task and make the necessary adjustments.

(c) To tell instantly whether his machinery produces the desired product or not by measuring the actual product and comparing it with the standard product.

(d) To know when machinery is to be altered or discarded, and another form substituted that will produce the standard product.

(e) To know whether the workmen are doing the thing that is

3. The Twelfth Yearbook of the National Society for the Study of Education, Part I, pp. 7-96, 1913.

expected of them, and thus detect good and poor work among the men.

(f) To know who needs help and who needs none.

(g) To know what one needs further training during service and what one would best be transferred to some other department or discharged.

(h) To have a basis for judging labor methods likewise.

A further illustration of the value of standards and tests in business enterprises is what Mr. Taylor, the efficiency engineer and originator of scientific management already referred to at the beginning of the chapter, calls the scientific selection of workers. Sometime ago, at the height of the bicycle industry, Mr. Taylor was asked to reorganize a large bicycle factory and make it as efficient as possible. Reference will be made here only to a group of 120 girls engaged in inspecting the steel balls used in the bearings, a work requiring clear vision, quickness and deftness of movement, and long-continued mental concentration. The girls were taken in as they came without any process of selection.

When Mr. Taylor started his work of reorganization, he immediately set up within the factory a psychological laboratory to study the native aptitudes and qualities of the girls. The girls were each tested in "perception-time" and "reaction-time". Only those quick in both perception and reaction were retained on the job. The same process of elimination was done in the mental concentration tests.

The result of the selection by tests was gain on every side.

Only 35 of the girls were employed, and these did the work with 60 per cent greater accuracy. The girls gained double wages, very much extended time for leisure, and improved health; the management saved 50 per cent of the original cost; and the public benefited from lower costs and the release of the remaining 85 girls for other branches of productive enterprises.⁴

(c) The Parallel Use of Standard Measurements in Educational Administration

Education is as much of a shaping process as the manufacture of rails or any other material production. It is the shaping of personality into desirable and useful forms. Education is one field of biological production, wherein the factor of growth plays an important part. But this acting alone will produce only an inferior output. Hence, the need of the shaping process and of standards by which to judge the results obtained and guide the administration in this work.

To illustrate the value of setting up standards in school products, the case of two city school systems, A and B, in the same place may be mentioned. The following show the state of affairs in the two systems:

GRADE	SUBJECT	SYSTEM A	SYSTEM B
VIII	Arithmetic (Combinations per minute)	35	105
VI	Handwriting (Ayres) - Speed (letters per minute)	58	115
	Quality	53	50

4. Ibid., pp. 63-64.

Commenting on this illustration, Prof. Bobbitt says:

"If so great variety is to be found among pupils of supposedly equal ability in these so-called standard subjects in matters the most simple, fundamental, and mechanical, we may reasonably expect that in the higher, more complicated matters of science, history, and the humanities, the variations are much greater."⁵

Setting up standards in school attainment will be of but little value if scales and methods to measure the educational product are not used for the purpose of determining whether the product rises to the standard set up. The value of such standards and scales accrues to the entire personnel of the school corporation.

(a) Value to the student

Forming the rank and file in the school, the students can know definitely what is expected of them, how much progress they are making, and to what degree of success they measure up to the standards of school achievement.

(b) Value to the teacher

The teacher, too, will know what is expected of her. In taking over a new class, she knows definitely what she is expected to do, what quality and amount of work she has to strive for in teaching this class. She can tell whether she is accomplishing much or little in her work and thereby rate her own self as good, medium, or poor. She will be able to know what help she needs and when, and discover the best methods for her work. She can apportion her time to different subjects on a scientific basis and determine how much she can accomplish within a given period of time. She will be a better judge of text-books and other school

5. Ibid., p. 14.

supplies. Lastly, seeing her own capacity as measured by the output she is producing, she can not be done any injustice, she can receive help in proportion to need, and recognition in proportion to merit.

(c) Value to the supervisor

By glancing over the records showing the test results in the achievement of pupils, the supervisor will be in a position to judge whether the teachers under him are securing the full results expected of them and whether they are handling the normal number of pupils. He can know which teacher to help and which teacher to encourage, which department needs the most aid and which the least. He can tell which teachers are improving by his suggestions and measure the amount of improvement.

With the aid of educational scales, the supervisor will have the best evidence of inefficiency against the weak teacher who can not be improved. And the problem of removing, transferring, or retiring such a teacher can be easily solved. The supervisor can know also when to relax his efforts and give no further help. He can judge the method of teaching his teachers are employing.

The supervisor can make various comparisons in different buildings. He can see the relative efficacy of different methods, text-books, appliances, and distributions of time by comparing the results secured. With scales of measurement, he can measure the efficiency of his principals. And under the circumstances, he will be giving promotion to those who by merit deserve it.

(d) Value to the superintendent

GRAPH SHOWING EFFICIENCY OF SUPERVISION

Enlarged from Bobbitt

SCORE - COMBINATIONS PER MINUTE - ADDITION

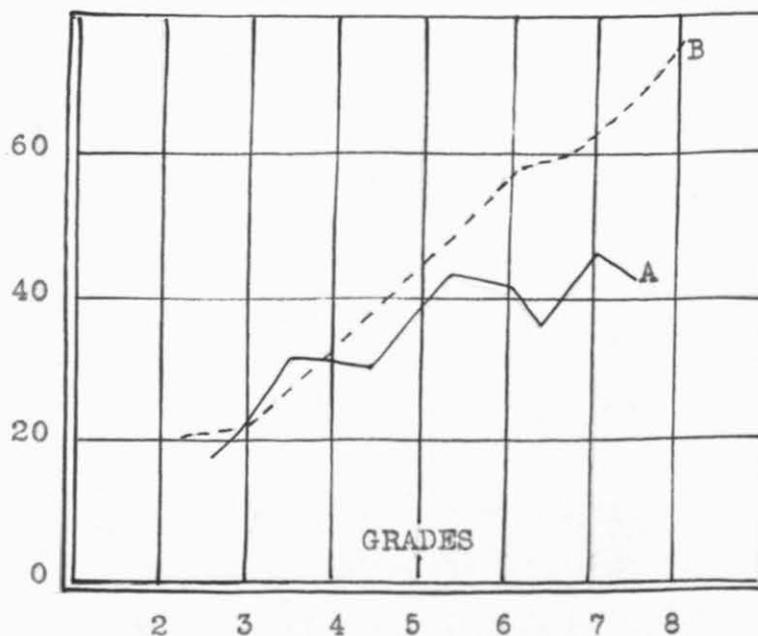


Fig. 1

A - Poor supervision. Much variation from grade to grade and a low final score. From an actual school.

B - Close supervision. Constant, steady growth from grade to grade and a high final product. From another school.

The superintendent will be able to locate the good, mediocre, and poor teachers in looking over the achievement results made by their classes, especially when these results are thrown into distribution tables and graphs. By a careful examination of the results in different buildings, he can tell the superior principal from the average and the inferior. He can tell then what help is needed, who need it, and in what particular line. And so he can send the right kind of expert from the central office to give the necessary assistance. He can measure not only the efficiency of his teachers, principals, and supervisors but also the efficacy of the methods, text-books, and appliances used.

Like the supervisor, the superintendent can make different comparisons. He can compare his different schools, or his own schools with others. He can, by so doing, see what task he has to accomplish to bring his schools up to standards. He may even go further. He can call the attention of the board to the size of the problem confronting the school system. In this way, he might succeed in securing more and better equipment, supplies, and teachers. And when it comes to appealing to the community at large, he can make himself understood by showing the results. Speaking in a language intelligible to the public, he can expect the support and co-operation he needs. He can urge on new improvements, projects, and innovations in the school system.⁶

(d) Summary

Scientific management has clearly demonstrated the value

6. Cubberley, E. P. Testing Results, Chapter XIX, pp. 326-328, in "Public School Administration", 1916, Houghton, Mifflin Company.

of standards and scales of measurements in business enterprises. The fundamental principles of scientific management resolve themselves into two: (1) setting up of standards and (2) employing scales of measurement to measure the product and determine whether it rises to these standards. The most efficient methods, the best trained workers, and the most useful appliances and devices are secured only after the setting up of standards and the using of scales of measurement.

The value of standards and scales is of no less importance in the field of education. Educational administration, like business administration, is dealing with the production of an output. As in the industrial concerns, the educational administrator needs to know what standards are required by the outside world and how to measure his results to determine whether these are rising to the standards set up. Hence, he feels the same necessity that the business manager does; namely, (1) the necessity for standards and (2) the necessity for scales of measurement. Standards and scales are of value not to the head administrator alone, but to his immediate and remote subordinates as well. The supervisor, the principal, the teacher, and even the student, profit by the use of standards and measurements. By the use of scales, the school system can be run with economy and efficiency, and the public can be appealed to easily for help, support, and co-operation.

CHAPTER IV

THE APPLICATION OF TESTS AND MEASUREMENTS TO AMERICAN EDUCATIONAL PROBLEMS¹

(a) Some Outstanding Problems in American School Administration

The problems of public education in the United States are so varied and numerous that only the most important ones and those that have some bearing on this dissertation are mentioned here.

The following classification is neither exhaustive nor exclusive:

- (1) Classification and Promotion of Pupils
- (2) Retardation and Elimination
- (3) Methods of Instruction and the Curriculum
- (4) Educational and Vocational Guidance
- (5) School Supervision and Organization

The specific and practical uses made of both the achievement and the endowment tests will be taken up in each of the above problems. No definite distinction is here made between the use of one test and that of the other for both have been used in most of these problems. Furthermore, it is pretty hard to draw a sharp boundary line between the two kinds of tests. As has been pointed out, achievement tests are also intelligence tests in many respects. And intelligence tests in their present form are not pure native endowment or potential ability tests. One test shades off

1. This discussion is based chiefly on: (a) Haggerty, M. E., Specific Uses of Measurement in the Solution of School Problems, Seventeenth Yearbook of the National Society for the Study of Education, Part II, Chapter III, pp. 25-40, 1918; and (b) Terman, L. M., Uses of Intelligence Tests in "Measurement of Intelligence", Part I, Chapter I, pp. 3-21, 1916, Houghton Mifflin Company.

into the other and when used together, achievement and intelligence tests supplement each other.

(b) The Problem of Grading and Promotion

(1) Educational Waste Due to Improper Classification and Promotion

One of the big sources of waste in Education is teaching what the bright pupils already know. Another is teaching the dull pupils what they can never know. According to Starch's findings:

"One third of the pupils waste time by being in classes in which they know practically all the material that is being covered in the recitation period and are able to perform all the tasks expected of them. Another third of the pupils waste time by being in classes in which they can grasp very little of the material and are able to perform very poorly, or not at all, the tasks expected of them. One pupil of every three is promoted too slowly and one pupil out of every three is promoted too rapidly. One pupil of every three could finish the eight grades in seven years or less, and one pupil in every ten could finish the eight grades in six years or less."²

Stated in unmistakable terms, such is the crux of the grading situation in the American school. A concrete illustration of this tremendous waste in education is shown in the amount of grade overlapping Starch has found in his experiment at Madison, Wisconsin. Tests were given in writing, reading, spelling, and arithmetic.³

a. Writing. - The tests were intended to measure both speed and quality. The speed was expressed in terms of the number of letters per minute, and the quality was graded with the Thorndike scale. The results showed that first grade pupils ranged all the

2. Starch, D. Standard Tests as Aids in the Classification and Promotion of Pupils. Fifteenth Yearbook of the National Society for the Study of Education, Part I, Chapter XIV, p. 143, 1916.

3. Ibid., pp. 143-148.

way from quality 4 to quality 9, and second grade pupils from quality 6.25 to quality 9. Such a wide range of abilities in each grade caused an enormous overlapping of the abilities of one grade over those of others. By actual computation, Starch found the amount to be as follows: (1) 32 per cent of the pupils in any grade equaled or exceeded the median of the next grade above; (2) the same per cent equaled or fell below the next grade below.

b. Reading. - The reading tests were intended to measure both speed and comprehension, but the two scores were combined in one. The results here were no less striking. There was as wide a range of abilities and as enormous an overlapping as in the case of writing. It was found that 31 per cent of the pupils in any given grade reached or exceeded the median of the next grade below.

c. Spelling. - The amount of overlapping was essentially the same as that of writing and reading. Twenty-three per cent of the pupils in any given grade reached or exceeded the median of the next grade above.

d. Arithmetic. - The Curtis scale, Series A, was used. Results told the same story. Thirty-two and five tenths per cent of the pupils in any given grade reached or exceeded the median of the next grade above.

The experiment was carried further to determine whether or not this wide range of ability and enormous overlapping would be very much reduced if each pupil's performance in all studies were combined. The combined tests showed practically the same result. Thirty-two and two tenths per cent of the pupils in any given grade reached or exceeded the standard of the next grade above; the

same per cent fell to or below the standard of the next grade below.

After making all allowances for any exaggeration Starch might have made in his study, one can still infer that: (1) the schools do not go by any definite, tangible standards and measure in any accurate way the actual abilities of children to determine whether they are up to standard or not; and (2) consequently, pupils who are one or more years ahead of or behind the grade in which they are placed (i. e. pupils who are misplaced in the grades) are overlooked.

Grade overlapping is no less marked in the case of mental development (or potential ability) than in that of school achievement (or performance) as shown by the overlapping of mental ages (the indices of mental level). The best illustration of this situation is given by Terman graphically in his study of typical California school systems.

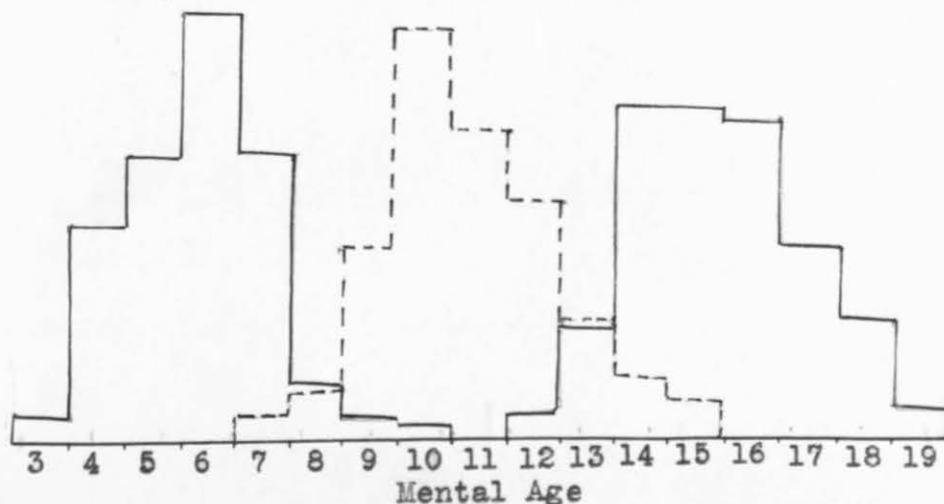


Fig. 2

OVERLAPPING IN THE MENTAL AGES OF CHILDREN IN THE
FIRST, FIFTH, AND NINTH GRADES
Enlarged from Terman

Figure 2 shows that: (1) the first grade children greatly overlap the fifth grade children; (2) the fifth grade children overlap the first year high school to about the same extent; (3) the brightest fifth grade child is above the median of the first year high school mental level; and (4) the brightest first grade child has reached the median for the fifth grade and nearly the mental level of the dullest first year high school child.⁴

Terman, commenting on improper classification, says:

"Notwithstanding the sifting which takes place at the end of each school year, the resulting classification of children has been so far from successful that, generally speaking, the lowest 25 per cent of pupils in any grade belong mentally in a lower grade and the highest 25 per cent in a higher grade. Only the middle half are classified approximately where they should be. Usually more than 15 per cent are at least two grades removed from the one in which they belong by mental age."⁵

The only remedy for this situation is to classify and promote pupils, not according to time or age, but according to performance and capacity. Such a method of classification and promotion would also give incentive to children to do their best. But in order to accomplish this task, definite standards must be set up for the various subjects, courses, grades, and schools. Then scales of measurement and tests must be used to ascertain whether these standards are being reached or not. The ordinary teacher's examination, besides involving the personal equation of the examiner, is not scientifically conducted and generally gives much worry and anxiety to the pupil. With mental ability as the best criterion for classification and promotion, this type of examination and

4. Terman, L. M. Individual Differences in "Intelligence of School Children", Chapter II, pp. 25-26, 1919, Houghton Mifflin Company.

5. Ibid., p. 27.

chronological age must give way to the achievement and the intelligence test in school grading.

(2) Changes in Classification and Promotion Due to the Use of Tests and Measurements

Dr. M. E. Haggerty reports several changes in the classification and promotion of pupils as indicated by the replies to the questionnaires he sent out to several school superintendents who had been making use of educational and mental measurements. He classifies these changes in a manner like this:⁶

a. Promotion and demotion of pupils improperly classified.

Such replies as these were received:

"We gave certain pupils double promotions." "We demoted and promoted pupils who did not fit the grade they were in."
"Bright pupils were put into relief by the tests and afterwards examined; a large number were promoted thus."

b. Classification and promotion by subjects. Under this heading are the following replies:

"Pupils were transferred in reading to grades for which the tests revealed they were fitted." "When pupils reached quality 10, Thorndike handwriting scale, in monthly tests they were promoted into advanced section, meeting three times a week." "Pupils classified in reading according to score in test: (1) those below Kansas standard drilled in thought interpretation; (2) those who equaled Kansas median given no extra attention; (3) those who tested a grade higher allowed to drop reading for a time and work on any study they were low in." "Promotion by subjects."

c. General class re-organization. Only one reply is quoted:

"I classified my school below seventh grade so pupils could make up work where they were weak and take advanced work where they were strong."

6. Haggerty, M. E. Specific Uses of Measurement in the Solution of School Problems. The Seventeenth Yearbook of the National Society for the Study of Education, Part II, Chapter III, p. 26, 1918.

Concluding on the value of tests in the proper classification and promotion of pupils, Haggerty says:

"Superior students can be detected and grouped together; mediocre students can be put with mediocre students, and weak students, instead of being submerged in the struggle to maintain standing, can receive the help they need. It would be difficult to overestimate the increase of efficiency that would come from the better adaptation of instruction in consequence of such classification."⁷

(c) The Problem of Retardation and Elimination

Retardation is generally defined as a state of over-ageness, a state in which a pupil is above the average age for his grade. Elimination is a state of dropping out on the part of the pupils for some reason, a condition that comes usually after retardation. Recently, retardation has been defined in terms of mental age instead of chronological age. Terman has taken this stand in his "Intelligence of School Children". According to this definition, it is the under-age pupil, not the over-age, that is a retardate.⁸

(1) Studies on the problem

Several studies of retardation and elimination have already been made in the United States. Ayres,⁹ Thorndike,¹⁰ Strayer,¹¹ and Bonner¹² have all gathered statistics to show the tremendous amount of retardation in American school systems. Terman, in reviewing all statistics from several hundred cities, says that from

7. Ibid., p. 28.

8. Terman, L. M. "Intelligence of School Children", Chapter V, pp. 72-74, Houghton Mifflin Company, 1919.

9. Ayres, L. P. "Laggards in Our Schools". 1909. Russel Sage Foundation; 236 pages.

10. Bulletin 4, 1907, U. S. Bureau of Education.

11. Strayer, G. D. Age and Grade Census of Schools and Colleges. Bulletin 451, 1911, U. S. Bureau of Education.

12. Bonner, H. R. Statistics of State School Systems (1917-1918), Bulletin 11, 1920; also Statistics of City School Systems (1917-18), Bulletin 24, 1920, U. S. Bureau of Education.

one third to one half of school children in the United States fail to progress through the grades at the expected rate, 10 to 15 per cent are retarded two years or more, and 5 to 8 per cent at least three years. As a result, about \$40,000,000 is spent annually for repeated instruction.¹³

(2) Causes of retardation

While other factors might play a part in retardation, - late entrance, irregular attendance, transfer to another school, physical defects, home and environmental conditions - the fact remains that under ordinary circumstances it is the more intelligent child that makes more rapid progress in the school and the mentally deficient that lags behind. After all, it is mental capacity that plays the leading rôle in school success. Of course, intellectual ability alone is not determinative of one's success in school life. It must be coupled with interest and industry. The capable pupil may fail because he lacks these other qualities and the mediocre pupil succeed because he possesses them. But with the same degree of application and effort, it is the more intelligent individual that wins out. And so it is with school pupils. The mentally weak lag behind and gradually drop out. Retardation and elimination are due in the main to lack of mental capacity.

(3) Application of tests to the problem

The Montclair (N. J.) schools' use of age-grade tables is the best illustration.¹⁴ Under the direction of Superintendent

13. Terman, L. M. "The Measurement of Intelligence", Chapter I, p. 3, 1916, Houghton, Mifflin Company.

14. Bliss, D. C. The Application of Standard Measurements

D. C. Bliss, the principals met to find out the amount of retardation in September, 1912. Figures showed it was 23 per cent. Steps were taken to discover the causes of this big percentage of retardation. Tests were administered to the children, the needs of the individual pupils as shown by the results of the tests were given more careful consideration and more flexibility was infused in the administration of the schools. Year after year, the superintendent and his corps of principals met to discuss the situation and its gradual improvement. The percentage of retardation decreased as follows:

1912	-	-	23.3 %
1913	-	-	18.3 %
1914	-	-	14.8 %
1915	-	-	13.7 %

The application of tests to the problem of retardation in the Montclair schools reduced the percentage almost 50 per cent.

(d) Methods of Instruction and the Curriculum

Haggerty's study of superintendents' replies to his questionnaires gives the following changes in the methods of teaching and the course of study:¹⁵

(1) Placing more value on certain subjects or parts of subjects and giving more time to these. Replies were received as follow:

to School Administration. The Fifteenth Yearbook of the National Society for the Study of Education, Part I, Chapter VI, page 71, 1916.

15. Haggerty, M. E. Specific Uses of Measurement in the Solution of School Problems. The Seventeenth Yearbook of the National Society for the Study of Education, Part II, Chapter III, pp. 31-37, 1918.

"More emphasis, all grades, on meanings of words and sentences." "Stressing legibility in writing." "Greater emphasis on fundamentals in arithmetic." "More emphasis on correct use of words in reading, less on definition." "Special emphasis given to those subjects where standard was low." "More silent reading."

(2) Increasing and specializing drill. Returns give these replies:

"Five-minute daily drill on fundamentals in arithmetic." "More time and attention given to drill in number combinations." "Dictation drills with attention on punctuation." "More intensive drill in grammar and punctuation." "More drill in spelling." "Courtis drill cards in arithmetic in two rooms; more oral drill in all." "Horace Mann method of spelling adopted." "Installed Palmer method of writing."

(3) Inventing special devices. The following replied thus:

"Made room charts showing individual's work (median, quartile, safety zone)." "Teachers used questions similar to Kelley test, and applied to geography and other subjects." "More instruction through interest of pupil." "Supervised study periods - all grades." "Three periods of supervised study added to school day: those who failed in one subject required to stay 1 period; in two, 2; etc." "Greater use of dictionary for meaning of words." "Tests devised to watch pupils' progress."

(4) Individualizing instruction. These replies come under this class:

"Individual attention; specifics devised for securing appreciation of good writing." "Individual help given slow pupils." "Methods adapted to ability of pupils." "Backward pupils discovered and given special attention."

Replies also showed the following:

(5) Changing the subject-matter taught or the text-book used.

(6) Setting up specific standards of achievement for different grades.

(7) Organizing special curricula for special classes.

(e) Educational and Vocational Guidance

(1) Educational guidance and standardized tests

Tests, more particularly the intelligence tests, are a great aid to the administrator of schools in placing the pupils in courses, classes, or schools where their mental ability fits them to be. The ultimate goal of the child is a life vocation. The only thing that mental and educational tests can give the administrator now is an approximate median mental level required for all vocations - the professions, trades, or common labor. Terman's findings and those of several others who worked with him show that law, medicine, engineering, teaching and the ministry require about the same amount of general intelligence, and that carpentry, masonry, plumbing, blacksmithing, etc., make about equal demands on general mental capacity.¹⁶ When the median has been determined for the different vocations from profession to unskilled labor, then it is the duty of the school to give such courses and subjects as will fit those pupils who, by the result of the tests, show they belong to this or that line of work. Mental testing, especially, has a very important rôle in educational guidance.

Educational guidance as a problem of school administration can not be overestimated in its importance. It is highly desirable that children are placed in the courses and given the subjects that lead to the vocations for which tests and measurements show them to be fitted. The junior high school advocates have realized the need of educational guidance more and more and have emphasized the inclusion of courses giving the child a chance to choose his vocation. But without the aid of mental tests no scientific placement of pupils can be brought about in the junior high school or

16. Terman, L. M. "Intelligence of School Children", Chapter XII, pp. 286-287, 1919, Houghton Mifflin Company.

any school preparatory to a vocation.

(2) Studies on the relation of tests to various vocations

Vocational guidance follows educational guidance. While educational guidance starts early in the child's career (about the fifth or sixth school year),¹⁷ vocational training comes later, before the child enters the outside world. Until definite mental requirements have been discovered for the different vocations, the field of vocational training in the school has to confine itself to general considerations. Tests as they are in their present form do not differentiate sharply between professions, between trades, or between classes of unskilled labor. But the first stage in vocational guidance and mental testing has already been attained—the differentiation of vocations into professional, trade, and common labor. The second stage of further differentiation in each of these vocations is bound to come in the future.

Several studies have been made to determine the level of intelligence for various vocations:

a. Terman's study of firemen and policemen. Applicants for positions in the fire and police departments of San José, California, were given mental and educational tests. The previous vocations of the examinees were classified into unskilled, semi-skilled, and skilled, and compared with the ranges of I. Q. and I. Q. averages. The following results were found:¹⁸

	Unskilled	Semi-skilled	Skilled
Range of I. Q.	63 to 89	74 to 96	84 to 112

17. Ibid., p. 269.

18. Terman, L. M. A Trial of Mental and Pedagogical Tests in a Civil Service Examination for Policemen and Firemen. *Journal of Applied Psychology*, 1917, pp. 17-29.

Average I. Q.	75.5	85.2	98.3
---------------	------	------	------

b. Flanders' study of express company employees.

This study shows that even in a group working on several tasks requiring about the same amount of intelligence, the range of mental ages is appreciably wide. Results from the 47 employees tested showed a range from 10 years (I. Q. 62) to 18-7 (I. Q. 116) and a median of 15-2 (I. Q. 95). This simply means tests are valuable in determining the level of mental ability in the different tasks that seem so similar as regards the amount of intelligence required.¹⁹

c. Knollin's study of business men. Thirty business men of limited schooling and moderate success were tested. The median mental age for the group was 16-2 (I. Q. 102). The lowest 25 per cent were below 15-0 (I. Q. 93.6); the highest 25 per cent above 17-2 (I. Q. 107).²⁰

d. Coover's and Downey's studies of college students. In sharp contrast to Flanders' and Knollin's studies of vocations requiring comparatively less intelligence, these two studies by Coover and Downey give results showing the different mental requirements for different vocations. Coover's study gives a median I. Q. of 113 and Downey's gives 108 for college upper classmen. These are certainly higher than 95 in Flanders' study and 102 in Knollin's.²¹

19. Flanders, J. K. Mental Tests of a Group of Employed Men, Showing Correlations with Estimates furnished by Employer. *Journal of Applied Psychology*, 1918, pp. 197-206.

20. Terman, L. M. "Intelligence of School Children", Chapter XII, pp. 278-279, 1919, Houghton Mifflin Company.

21. *Ibid.*, pp. 279-282. Also Downey, J. E., The Stanford Adult Intelligence Tests, *Journal of Delinquency*, 1917, pp. 144-155.

(f) School Supervision and Organization

Much has been said in previous sections of the general value of standardized tests to educational administration and supervision. An attempt was made early in this study to enumerate the specific uses of tests and measurements to the administrator. But the practical working of the tests as they affect educational administration has not been touched yet. In giving a few instances of the practical use made of tests in school supervision and organization, reference will again be made to Haggerty's questionnaires.

According to this study, some of the changes in school organization and administration as a result of using tests and measurements are:²²

(1) Changing size of, and sectioning, classes. The following replies were reported:

"Smaller classes in arithmetic." "More teachers in arithmetic." "The services of additional teachers demanded for defectives in industrial training." "Enlargement of class for defectives." "An advanced and a special section made in writing on the basis of errors."

(2) Organizing special classes and schools for backward and gifted children. These replies were noted:

"Opened special room for backward pupils." "Organized special corrective work." "An initial attempt to develop an elementary industrial school for pupils shown by the tests to be unfitted for the regular work."

In this connection, the Montclair experiment with subnormal and precocious children might be mentioned as another practical ap-

22. Haggerty, M. E. The Specific Uses of Measurement in the Solution of School Problems. The Seventeenth Yearbook of the National Society for the Study of Education, Part II, Chapter III, pp. 29-31, 37-39.

plication of tests and measurements to the problem of special class organization. Tests were given to a class of subnormal children in arithmetic and handwriting. The progress of each child was recorded from September to June. Results showed the backward pupils made steady progress in handwriting but a great fluctuation was noticeable in the case of arithmetic. The second part of the experiment deals with gifted children. A special class of these children, who were recently promoted to the seventh grade after making three years' work in two, were tested in fractions, English, spelling, writing, fundamentals, and composition, together with other seventh grade classes. Results showed that the special class scored much higher than the others in all the examinations except handwriting and spelling. It is evident from this experiment that the gifted child succeeds more in the abstract processes and improves but little in the mechanical tasks, while the backward child benefits only from the latter and fails to grasp abstract ideas. By the proper utilization of the results of experiments like this, the Montclair school system has been able to cope with retardation and other educational problems.²³

(3) Departmentalizing the grades. Haggerty's study already referred to gives the following replies:²⁴

"Departmentalization of sixth, seventh, and eighth grades."
"Department teaching."

(4) Appointing supervisors, supervising principals, and directors of efficiency and research bureaus. The following come

23. Bliss, D. C. The Application of Standard Measurements to School Administration. The Fifteenth Yearbook of the National Society for the Study of Education, Part I, Chapter VI, pp. 73-78.

24. Vide note 22, p. 48.

under this:

"Position of supervising principal for primary grades created." "Appointed a director for a newly organized Bureau of Research." "Plan to hire a trained supervisor for writing next year."

Besides these initial steps taken to improve supervision, several changes were also indicated by the replies as a result of re-testing to ascertain the amount of improvement. These replies are typical:

a. In arithmetic. "Graph showed greater improvement in grade in one month, than first graph showed from class to class (five months' work)." "Have used drill and are now above the Courtis standard." "The weaker pupils do better work. More pupils brought up to required standard. Pupils more accurate." "Plateau disappeared which had existed from Grade VI to VIII; curve for both attempts and rights in all operations shows gradual development to Grade VIII."

b. In reading. "Improvement in median score from September 7 to January 23; Grade VII, 5.3, Grade VIII, 5.6."

c. In spelling. "Pupils average a grade higher." "More uniformity of grades and children know where they stand."

d. In writing. "Better quality. Both quality and speed nearer average for grade."

(5) Training teachers for increased efficiency. This study not only shows considerable improvement in the school achievements, in the pupils, and the school as a whole, but it also gives a further change in the teaching corps, a change for increased efficiency in the service. Replies from the superintendents who had used tests and measurements showed such proposed improvements in the teaching service as these: a. Introduction of teachers' training; b. summer-school attendance for certain teachers; c. visit to demonstration or model classes; d. teachers' shifting from grade to grade; and e. changing principals.

(g) Summary

Tests and measurements have given the American public school administrator: (1) a proper basis for pupils' classification and promotion; (2) an instrument to diagnose the amount of retardation and elimination and a remedy for its solution; (3) a standard by which to judge the efficacy of teaching methods and the arrangement of the program of studies; (4) a general criterion for educational and vocational guidance; and (5) a most powerful agency in bringing about standard school achievements, better supervision, establishment of research bureaus, better trained teachers, and an efficiently organized school system.

The studies made by Haggerty, Terman, Bliss, Starch, and a host of other investigators show very clearly the valuable uses to which standardized tests have been put in the American school system. Tests have shown the wide range of abilities in school children and the consequent enormous amount of overlapping in the grades, when reckoned in terms of either performance or native capacity. They have shown the tremendous percentage of retardation and elimination as a result of improper classification. They have led to more efficient methods of instructions and more differentiated courses of study to meet individual needs in the school population. They are giving a fairly reliable basis of prediction for the child's success in school and in life. They have brought about changes in the teaching staff, supervision, and general organization of the school system, changes making for efficiency and economy. In short, tests and measurements have been playing a significant part in American public school administration and have convincingly demonstrated their practical value in helping solve

the many and varied problems of the school.

CHAPTER V
THE APPLICABILITY OF STANDARDIZED TESTS TO PHILIPPINE
PUBLIC SCHOOL CONDITIONS

(a) The Evolution and Administration of the Philippine Public School System

(1) Brief history of Philippine public education

The history of Philippine public education can probably be best understood if a study is made of its background, the early days of Malayan culture and the three centuries and a half of Spanish occupation.¹ But to trace such a development would entail a considerable amount of space devoted to historical facts that may have little or no bearing on our subject, and consequently may carry our discussion far from the topic at hand. We shall, therefore, begin with the transplanting of the American public school system, go on to describe some of the subsequent changes effected to suit local needs and conditions, and finally consider the Philippine public school in the light of modern conceptions of public education.

With the establishment of primary schools by the army,² two outstanding problems confronted the early school authorities; namely, the training of a Filipino teaching staff to carry on the instruction begun by the Americans and the publication of textbooks suited to Philippine needs. The first problem was met first, by the inclusion of normal training in the upper grades taught by

1. Vide "Recopilacion de las Leyes de Indias" (ley V., tit. XIII., lib. I.). Also Foreman, J., "The Philippine Islands", Chapter XI, pp. 190-196, Charles Scribner's Sons, 1899.

2. Elliott, C. B. "The Philippines, to the End of the Military Régime", Chapter XVII, pp. 485-486, The Bobbs-Merrill Co., 1916.

the Americans; secondly, by the erection of normal schools throughout the Islands; and finally, by the establishment of the College of Education, University of the Philippines. Not only has this task of training teachers proved to be successful, but the process (popularly known as Filipinization) has gone on so rapidly that of the 17,000 teachers in the Philippine public schools today, only about 350 are Americans.³ The second problem was readily met by a gradual process of modification in the text-books used and the subject-matter taught. It was discovered from the outset that the Filipino child, with a different interest, owing to different surroundings, did not delight much in reading about Jack Frost, the squirrel, or the apple; nor did he take any deep interest in Indian or Norse stories. Steps were immediately taken to replace snow and frost with rain, the squirrel with the monkey, and the apple with the banana. Today the Filipino child reads not so much of old Indian and Norse legends as of his native folklore stories, not so much of Jack and John as of Juan and Pedro.

(2) Philippine public school administration

The administration of the Philippine public schools combines the democratic control in the American system and the centralized authority in the Prussian and the French. At the head of all the schools is the Director of Education with his two assistants. He is directly responsible to the Secretary of Public Instruction, who is at the same time Vice-Governor of the Islands and a member of the Council of State, a body composed of the Governor-General, the six department secretaries, the President of the

3. From Elementary and Secondary Educational Statistics, Philippine Islands, March, 1920.

Senate, and the Speaker of the House. The director is appointed by the Governor-General, with the advice and consent of the Senate, on the recommendation of the Secretary of Public Instruction and the Council of State. Wielding all authority vested in his office, the director is held responsible for the efficiency of the school system by the direct representatives of the people.

Provincial, municipal, and other local control is vested in fifty-two division superintendents, who are appointed by and are responsible to the director. These division heads appoint all district supervisors, elementary and secondary principals, and all teachers on the recommendation of their respective departments. The efficiency of the teaching personnel is maintained by a series of civil service examinations, normal institutes, and summer courses offered by the University of the Philippines. Teachers who are neither normal school graduates nor holders of teachers' certificates may qualify for permanent appointment only after passing the junior teachers' examination; those who are already regular appointees in the service, seeking promotion in rank or salary, may take the senior teachers' examination and become eligible upon making a satisfactory rating; while prospective division superintendents are required to pass the assistant examination, the highest given by the Bureau of Civil Service, open only to clerical and educational executives. At the city and provincial normal institutes, so-called model teachers conduct demonstration classes, and prominent educators give lectures and lead discussions on educational problems. During the summer, the College of Education at Manila and Baguio (where annual teachers' conventions are held)

offers very profitable professional courses to those who have the time and energy to spare.

The four main functions of Philippine public education are assigned to separate departments from the central office down the line. The academic phase of the work is headed by the chief of the academic division, Bureau of Education, and handled by selected academic supervisors in the provinces and municipalities. The industrial division of the main office takes charge of the rather extensive program of industrial and vocational training in the schools from the first grade through the vocational secondary schools. Local control is left in the hands of special industrial supervisors. The third and fourth functions, those of physical training and social recreation, are handled by the physical and playground director, who is in Manila, and leaves local administration to a corps of athletic coaches and playground instructors. This department of physical education and social recreation takes care of all athletic and playground activities.

School funds in the Philippines are taken chiefly from insular appropriations voted by the Legislature after the preparation of the annual budget by the Council of State. Funds also come from voluntary contributions and donations by private individuals or corporations. And not infrequently governmental authorities (provincial boards and municipal councils) vote special funds for the erection of new buildings or the establishment of new schools in addition to insular aid from the Legislature. After funds have been secured, the business and financial phases of the work in the Bureau of Education is handled by the property and supply, and

and accounting departments respectively.

(b) Measurements in Solving Philippine Public School Problems

In the light of the foregoing discussion, the most important problems of Philippine public school administration may be grouped as (1) educational and (2) business.

(1) Educational problems

With such a close resemblance between the American and the Philippine system of education, the Philippine administrator meets practically the same problems as the American. As these have already been treated in detail in previous sections of this study, only a brief summary is here necessary:

a. Classification and promotion of pupils. Like the American, the Filipino child is classified on the basis of performance in the ordinary teachers' examination.

b. Retardation and elimination. The enormous mortality in the first year high school invites the attention of the Philippine administrator to a thorough study of retardation and elimination.

c. Methods of teaching and the curriculum. The efficacy of the so-called model methods from the city systems and the selection of subject-matter for the courses of study deserve likewise careful consideration on the part of the Philippine school authorities.

d. Educational and vocational guidance. Nowhere will the need for a scientific procedure in these lines be more felt than in a school system like that of the Philippines where extensive industrial training is given in all grades.

e. School supervision and organization. If America has felt the necessity for better supervision and organization in the schools, there is every reason to believe that the highly centralized system in the Philippines will readily show such defects in instruction, supervision, and administration and give a very important task for the administrator to solve.

(2) Business problems.

Being a centralized machinery, the Philippine public school system owes its existence to the people through their representatives in the Legislature. An appeal that is made to them for school funds will be listened to only when this is made in terms plain, concrete, and intelligible. The people express their voice through their chosen representatives in the Council of State, Senate, and House of Representatives. The Secretary of Public Instruction is the only instrument of the schools in asking for the necessary appropriation to maintain and improve them. He is in the same position as the superintendent of schools in America, fighting to secure the necessary amount for his schools.

(3) The Rôle of Tests

We have seen the rôle of standardized tests in helping solve the problems mentioned above in American education. In several systems, pupils have been classified and promoted according to mental ability. Retardation has been greatly reduced. Methods of instruction and courses have been much improved. Educational and vocational guidance have gained greater impetus. Educational supervision and administration have become more efficient.

What we can expect of tests and measurements in America we have every reason to expect in the Philippines. With the use

of tests and measurements, the improper classification of pupils can be discovered and proper grading installed. The causes of the enormous mortality in the first year high school can be studied along with the general problem of retardation and elimination. The most efficient methods of instruction and the most highly differentiated courses of study can be employed. The Philippines' pride of giving an extensive vocational training can be placed on a scientific foundation and bids fair to stand any modern change in the field of education. Then the task of general supervision and organization will be very much lightened when the school authorities have some tangible standards to go by in judging school efficiency. Finally, the help and support of the people through legislative action, and voluntary contributions can be easily secured if the educational representatives in the government, the Secretary of Public Instruction and the Director of Education, can present their estimates of school needs and expenditures in terms of school achievement and pupils' progress in the grades by using the results obtained from standardized tests. Granted that the use of standardized tests will help Philippine public school officials solve their problems, these questions arise: Will the tests as arranged and standardized in their present form apply to Philippine conditions? Will they be a fair test of the Filipino pupils' school achievement and native capacity? Are any changes needed?

(c) Studies on the Applicability of Tests to Filipinos

While Philippine public school administrators have always followed the course of educational events in America and tried to

borrow as much as possible of American educational thought and procedure, very little has been done so far in the field of measurements. With the limited means at his disposal, the writer was able to carry on only one of these investigations, but through the cooperation of the Second Assistant Director of Education, Manila, and Ohio State University, Columbus, Ohio, he is able to present here two additional studies on the applicability of tests and measurements to Filipinos.

(1) Study on Filipino students in Minnesota.

About the middle of the winter quarter (1920-21), twenty-two Filipino students residing in Minneapolis and St. Paul were given the Haggerty Intelligence Examination, Delta 2. The following scores were made:

TABLE I

SCORE	FREQUENCY	SCORE	FREQUENCY
24	- 1	99	- 1
61	- 1	108	- 1
70	- 1	111	- 1
71	- 1	114	- 1
85	- 1	119	- 1
89	- 1	125	- 1
90	- 1	127	- 2
91	- 1	132	- 1
93	- 1	137	- 1
96	- 1	141	- 1
97	- 1		

The median for the group is 98.3. The range is from 24 to 141, an unusually wide one for a group of this size. The range of ages (chronological) is from 17 to 27, that of school attainments from first year high school to second year college. Of the twenty-two tested, there are five university students, eight private college,⁴ four high school (Minneapolis East High), and

4. Hamline University, St. Paul, and Minnesota College of

five are not in school.

The medians for the four groups are:

- (1) University - 125 (Range - 97 to 137)
- (2) Private College 100 (Range - 24 to 141)
- (3) East High - 100 (Range - 85 to 127)
- (4) Not in School 80 (Range - 61 to 96)

Of the 5 university students, 4 are above the entire group median; of the eight private college, 5 are above this median; of the 4 East High, 2 are above; while none of those out of school reaches or exceeds the group median.

In order to determine the relative difficulty of the different items of the test, the percentage of score made by the students to the total possible score was computed for each exercise or item. The following summary gives the medians of percentages in the six exercises:

Exercise 1 - - Sentence Reading - -	Median	62 %
Exercise 2 - - Arithmetical Problems	"	58.3 %
Exercise 3 - - Picture Completion -	"	53.7 %
Exercise 4 - - Synonym-Antonym - -	"	57.5 %
Exercise 5 - - Practical Judgment -	"	42.8 %
Exercise 6 - - Information - - -	"	71.6 %

It is evident from this study that: (1) the Delta 2 intelligence examination retains its discriminative capacity by dividing the group of 22 students into mental abilities ranging from that shown by score 24 to that shown by score 141; (2) those who generally find some way of continuing their studies score much

higher (medians 125 and 100) than those out of school (median 80) - all of these students being either entirely or partly self-supporting; (3) English as a language practically foreign to these students plays but an insignificant part in dividing the groups; and (4) mental ability, therefore, is the most important factor in differentiating these students from one another. Of the six exercises, the one that demands the least amount of use of language (picture completion) fell below all the rest except Exercise 5 in the percentages (score made to total possible score).

(2) Study on Filipino students in Ohio

Twelve students attending Ohio State University were given the Army Alpha intelligence examination and nineteen others the Ohio State University group test in the school year (1920-21).

a. Results from the twelve students in the Army test are recorded as follows:

TABLE II

SCORE		FREQUENCY	SCORE		FREQUENCY
34	-	2	78	-	2
46	-	1	95	-	1
50	-	1	96	-	1
59	-	1	122	-	1
69	-	1	123	-	1

The median is 70, range 34 to 123. Range of school attainments is from first year to third year college. Of the twelve tested, two are taking veterinary medicine, five are in the Arts College, three are taking engineering, and two education. It is interesting to note that the two highest scores were made by the latter group, 122 by a freshman and 123 by a junior. Of the two veterinary students, none reached the median of the entire group; of the five arts two were above; of the three engineering two got

a little above; and the two education got the highest.

The following medians of percentages of score made to total possible score have been computed to compare the relative difficulty of the exercises:

Exercise 1	- - Directions	- - - Median	50 %
Exercise 2	- - Arithmetic	- - - "	48 %
Exercise 3	- - Practical Judgment	"	40 %
Exercise 4	- - Synonym-Antonym	- "	30 %
Exercise 5	- - Disarranged Sentences	"	22.5 %
Exercise 6	- - Number Series Completion		5.4 %
Exercise 7	- - Analogies	- - - "	35 %
Exercise 8	- - Information	- - - "	50 %

The above study shows that: (1) even in a more homogeneous group as this, the Army intelligence examination gives a wide range of variation in the scores; (2) that those higher in college attainment are not necessarily higher in intelligence for one freshman scored 122, one of the two highest, much above four sophomores and one junior; (3) that again the language element fails to play a significant part in the scores, for Number Series Completion (requiring the least amount of English), went down to 5.4 %, below all the rest; and (4) therefore, native endowment is the principal factor, and not schooling or knowledge of English.

b. Results from the Ohio State University test. The nineteen students scored as follows:

TABLE III

SCORE	FREQUENCY	SCORE	FREQUENCY
32	- 1	56	- 4
34	- 1	59	- 1

SCORE	FREQUENCY	SCORE	FREQUENCY
38	- 1	65	- 1
40	- 1	69	- 1
42	- 1	70	- 1
45	- 1	74	- 1
47	- 1	86	- 1
51	- 1	109	- 1

The median is 56, range is from 22 to 109. School attainments range from first year to fourth year college. Of the nineteen students tested, 6 are engineering students, 1 medicine, 5 veterinary, 3 arts, 3 agriculture, and 1 commerce and journalism. The highest score was made by a sophomore in the Arts College. Of the 6 engineering students, 4 reached or exceeded the group median; the one medical student exceeded it by 3 points; of the 5 veterinary, 4 reached or exceeded it; of the 3 arts, 2; of the 3 agriculture, none; the commerce student was 18 points below.

A similar computation of the relative difficulty of the different exercises in the test showed Exercise 5 on Number Series Completion the lowest in per cent of score made to total possible score. The items are: a. Arithmetic (42.1 %), b. Proverbs (18.7 %), c. Synonym-Antonym (38.3 %), d. Disarranged Sentences (27 %), e. Number Series Completion (6.8 %), f. Reasoning (49 %), g. Analogies (12.5 %), and h. Information (35.8 %). This study merely supports the conclusions drawn from the one immediately preceding.

(3) Study on Filipino Teachers in the Philippines

Two groups of Filipino teachers were given the Otis Advanced test in the summer of 1919 at Baguio, the summer capital, and in four provinces.

a. Results from the provincial group give the following scores:

TABLE IV

SCORE	FREQUENCY		SCORE	FREQUENCY		SCORE	FREQUENCY		SCORE	FREQUENCY			
	M	F		M	F		M	F		M	F		
3	1		49	9	7	84	12	11	119	7	2		
10	1		50	13	12	85	15	6	120	2			
13		1	51	16	10	86	9	11	121	3	1		
15	1	1	52	7	8	87	15	12	122	9	3		
16	1		53	7	11	88	12	2	123	2	1		
18	1	1	54	9	13	89	15	8	124	2	2		
20		2	55	11	10	90	13	9	125	2	2		
21	2	4	56	14	7	91	8	5	126	4			
22	2	1	57	12	10	92	8	7	127	2	1		
23	1	2	58	13	14	93	10	6	128	3			
24	1	2	59	11	11	94	13	5	129		2		
25	1		60	16	13	95	16	9	130	1	5		
26	3	1	61	14	14	96	8	3	131	4	1		
27	3	2	62	14	20	97	15	4	132	2			
28	1	3	63	11	9	98	9	5	133	1	1		
29	2	4	64	16	13	99	9	11	134	3	1		
30	3	4	65	12	11	100	9	6	136	2			
31	3	8	66	18	10	101	15	9	137	5			
32	2	4	67	16	13	102	7	4	138	1			
33	4	4	68	17	18	103	2	3	139	2			
34	7	4	69	12	11	104	6	3	140	1			
35	7	2	70	14	10	105	8	6	141	1			
36	3	7	71	21	16	106	7	4	142	2			
37	3	6	72	16	10	107	7	3	144		1		
38	7	3	73	16	2	108	7	6	147	2			
39	3	4	74	18	23	109	6	5	149	1	1		
40	5	5	75	26	8	110	7	2	150	1			
41	5	5	76	14	7	111	4	2	151		1		
42	14	6	77	15	11	112	3	1	153	1	1		
43	6	6	78	16	8	113	4	2	154		1		
44	11	7	79	10	12	114	6	3	156	1			
45	3	8	80	18	11	115	5	3	161	1			
46	7	17	81	15	12	116	8		163		1		
47	8	9	82	18	14	117	6	2	167	1			
48	10	8	83	9	10	118	1	5	171		1		
										Total		1000	752

The average score for the men is 77, for the women, 70.⁵

Figure 3 gives the data in graph form.

5. Data from the Second Assistant Director of Education, Manila, collected in terms of averages, the writer has employed averages throughout this study in order to avoid confusion.

PERCENTILE DISTRIBUTION OF SCORES FOR MEN AND
WOMEN - PROVINCIAL GROUP

Enlarged from van Heyningen Hartendorp

Percentage

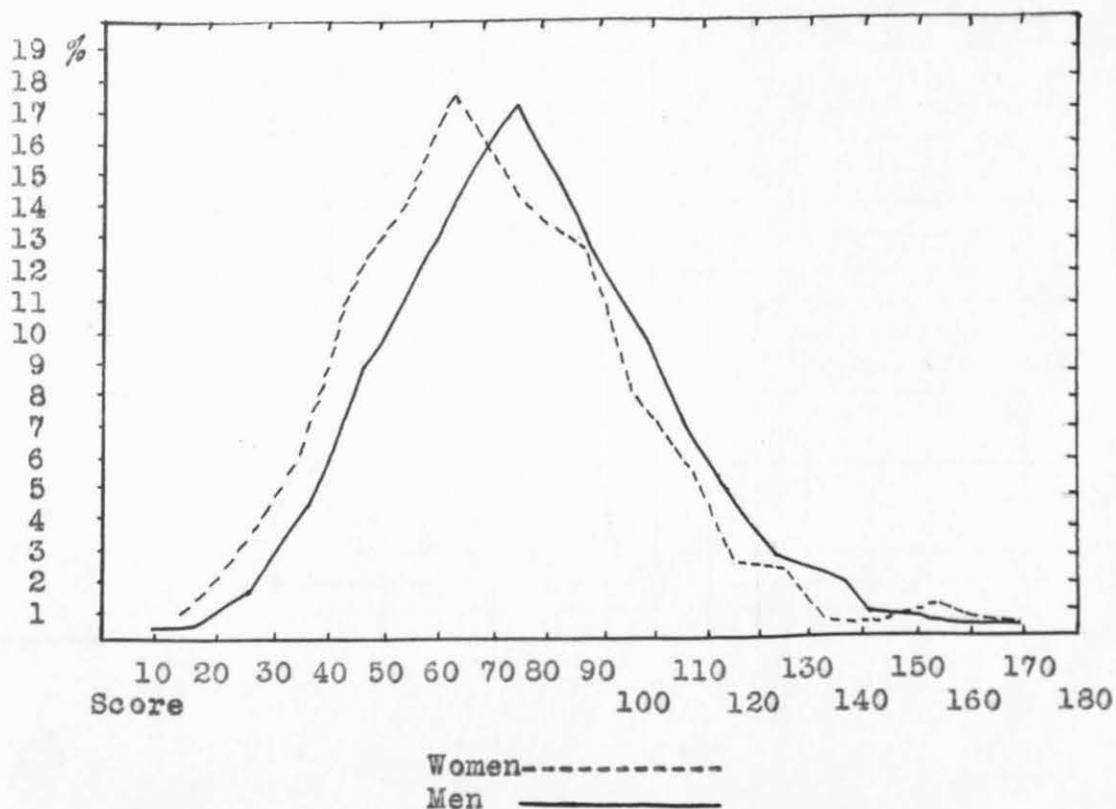


Fig. 3

The average scores by provinces for men and women separately is given in Table V.

TABLE V

Men

Province	NO.	1	2	3	4	5	6	7	8	9	10	Total
Rizal	152	8.4	9.7	10.8	3.7	9.9	6.4	7.4	7.3	8.2	11.0	82.8
Tayabas	366	7.9	11.4	8.4	4.7	8.7	6.0	5.7	6.1	8.8	9.5	77.2
Laguna	257	7.6	9.6	8.2	4.0	9.3	5.9	7.2	6.6	6.8	10.0	75.2
Batangas	225	8.1	9.1	8.3	3.3	10.3	6.3	7.1	6.8	6.4	9.4	75.1
Total	1000	7.9	10.2	8.7	4.0	9.4	6.1	6.7	6.6	7.6	9.8	77.0

Province	No.	Women										Total
		1	2	3	4	5	6	7	8	9	10	
Tayabas	176	7.4	11.3	9.0	4.3	7.3	5.5	5.1	5.3	8.6	9.4	73.2
Laguna	209	7.4	9.4	8.0	3.5	7.9	5.6	6.0	5.9	7.7	9.9	71.3
Rizal	203	7.5	8.0	8.7	2.3	8.4	5.6	6.7	7.1	6.7	9.3	70.3
Batangas	164	7.5	8.4	6.9	2.6	8.4	5.6	6.5	6.6	5.8	10.0	68.3
Total	752	7.5	9.3	8.2	3.2	8.0	5.6	6.1	6.2	7.2	9.6	70.9

Table VI gives the average scores by educational attainments.

TABLE VI

	NUMBER		SCORE	
	(men)		(women)	
Below seven grade	50	56.9	53	53.4
Intermediate graduates ⁶	509	70.5	498	69.1
First year H. S. and N. S.	196	73.5	125	70.8
Second year H. S. and N. S.	83	85.6	39	85.5
Third year H. S. and N. S.	38	93.4	6	95.0
Fourth year and graduate H. S.	83	103.0	23	109.3
Fourth year and graduate N. S.	28	115.6	88	113.6
College graduates and undergraduates	13	110.5		
	1000		752	

b. Results from the city (Baguio) group are given in the following table:

TABLE VII

SCORE	FREQUENCY										
	M	F		M	F		M	F		M	F
54	1		58	1		61	1		62		1

6. In the Philippines, there are only seven elementary grades - the first four in the primary school and the last three in the intermediate. An intermediate graduate has completed the seventh grade.

SCORE	FREQUENCY											
	M	F		M	F		M	F		M	F	
71		1	102	6	1	124	3	1	146	2	1	
75	1		103	2		125	6		147	3		
77		1	104	1		126	2		148	2		
78	1	2	105	2		127	1	1	149	1		
79		2	106	1		128	4	2	150	1		
80	1		107	3		129	5	1	152	2		
82	1	1	108	2		130	1		153	1		
83	1		109	2		131	6	3	154	2		
84	1		110	3		132	2	2	155	1		
85	1	2	111	2	3	133	4		156	1		
86		3	112	4		134	2		157	1	1	
88		1	113	6		135	1		158	2	1	
90	1	1	114	4	1	136		4	161	1		
91		1	115	2	1	137	2	1	162		1	
92	1		116	5	1	138	2		163	1		
93	4		117	2	3	139	2	2	164	1		
95	1		118	2		140	1		166	2		
96	5	1	119	1	2	141	2	1	171	1	1	
97	1		120	3		142	2	1	183	1		
99	2	1	121	3		143	1	1	189	1		
100	1		122	2	2	144	2	1				
101	1		123	3		145	1					
										Total	166	59

The average score for the men is 122, for the women 115.

Figure 4 gives the same data graphically.

PERCENTILE DISTRIBUTION. BAGUIO GROUP

From van Heyningen Hartendorp

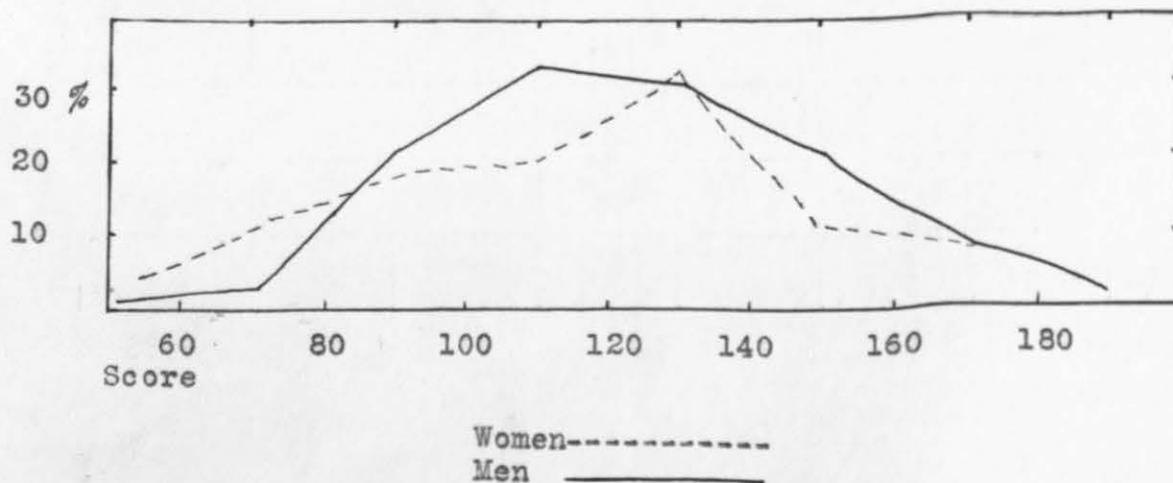


Fig. 4

Table VIII gives the average scores by educational attainments.

TABLE VIII

	NUMBER	SCORE	NUMBER	SCORE
	(men)		(women)	
First year H. S. and N. S.	46	112.1	18	95.5
Second year H. S. and N. S.	13	116.0	3	126.3
Third year H. S. and N. S.	11	116.1	7	115.0
Fourth year and graduate N. S.	47	128.2	22	125.9
Fourth year and graduate H. S.	16	129.7	4	106.7
College - graduates and under-graduates	33	129.9	5	136.8
	<hr/>		<hr/>	
	166		59	

Figure 5 shows the average scores in each test for both groups in graphical form. (See page 70).

From the above data, we find: (1) that the Baguio or city group, being a highly selected group, scores much higher than the provincial group; (2) that education affects to some extent the scores made as shown by the tables on educational attainments. But it is very possible that those who could not go on with their studies would not have succeeded much even if they did continue; (3) that education makes less difference than native endowment in the scores obtained, for scores show that first year high school and normal men in the provincial group scored 73.5, while the same group among the Baguio teachers scored 112.1; and (4) that the large gap - 15 points - between the score of teachers who are intermediate graduates (70.5) and those who are second year high school students (85.6) throws some light on the enormous mortality⁷ in the first year high school class of the Philippines.

7. Philippine Journal of Education, Vol. III, No. 3, p. 59, September, 1920.

PROFILE GRAPH SHOWING AVERAGE SCORES IN EACH TEST

Enlarged from van Heyningen Hartendorp

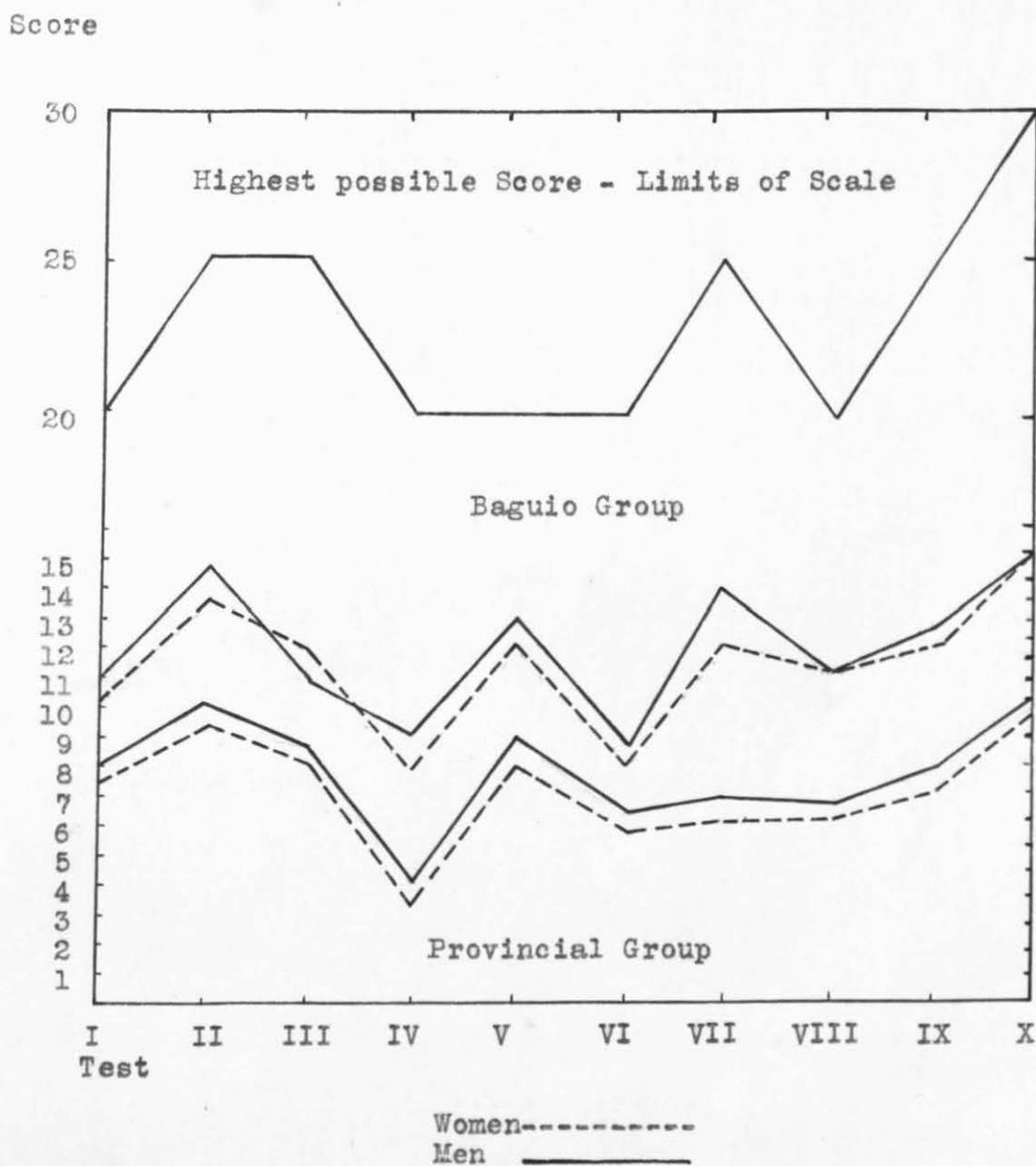


Fig. 5

(4) General conclusions

The studies given above enable us to draw some important conclusions on the applicability of tests to Filipinos. Among these are:

a. The intelligence test as devised and standardized in America retains its discriminative value when applied to the Philippines.

b. Education or schooling plays but a small part in the Filipino's general ability to perform the intelligence test.

c. The influence of language, which at first thought would seem to greatly handicap the Filipino, is practically negligible in the scores on the mental test.

d. The fact that the average or median for Filipinos is lower than the standards given for Americans must be explained not so much in terms of schooling or knowledge of English as in other factors that may come in.

e. As tests for the Philippines alone, tests in their present form give practical benefits as an aid to public school administrators; but as the basis of comparison with American norms, they show some discrepancy between the results in the two countries.

f. Changes, then, must necessarily be, not along lines pertaining to the use of English, but along the informational, subject-matter, and other phases of the tests.

(d) Summary

The Philippine public school system has germinated from the American idea of universal education, carried over and transplant-

ed on the soil of early Malayan culture and Spanish civilization. While partaking of the nature of European centralized educational administration, it does not differ very materially from the American system. Like American public education, the objective of Philippine educational theory and practice is the all-round development of the individual with a view to discovering his special innate faculties and giving him intensive training in his chosen field. As in the American school, English is the only language of instruction. And with the established policy of making the language of democracy the ultimate national medium of thought, writing, and speech, coupled with the general tendency at the present time to use English not only in commercial and educational circles but also in political and social gatherings as well, the rise of an English-speaking people on the other side of the Pacific is inevitable.

Modeled after American theories and practices, Philippine public education has presented the same problems to the school administrator. And so following the example of American educational authorities, public school officials in the Islands have to make use of all available means to cope with these problems. The American administrator has found tests and measurements an invaluable aid to the running of his school system. There is every reason to believe that the Philippine administrator will avail himself of the wonderful opportunities of this great movement in education.

Three important, though merely preliminary, studies on the applicability of tests to Filipinos have shown that: (1) the tests as arranged and standardized in America retain their discriminative

capacity when applied to Filipinos, thus meeting one very valuable criterion for any satisfactory mental measurements; (2) the influence of schooling and knowledge of English is much less than is generally thought to be as shown by comparison of individuals of different educational attainments and of the relative difficulty of the various items in the test; and (3) for use among Filipino school children alone, without making an attempt to draw comparisons with results obtained in America, the tests as they stand serve a very good purpose in Philippine public school administration. Changes, if ever made, to adapt tests to Philippine needs and conditions and do justice to the Filipino child when it comes to making comparisons with American standards, must be made carefully and thoroughly without invalidating the usefulness of the tests by substituting tasks on information and subject-matter of about the same difficulty to the Filipino child as the original items are to the American. Only by additional studies on the applicability of tests and measurements to Philippine educational conditions can the actual need for such changes be discovered and necessary modification brought about. Tests and measurements, then, besides assisting the public school administrator in the Philippines in the immediate solution of educational problems, lead to a wide field of scientific investigation, far-reaching in its importance, not only to Philippine public education but to the whole realm of educational science.

CHAPTER VI
SUMMARY AND CONCLUSIONS

(a) The Testing Movement in America and its Educational Significance

Education has never been considered much of a science and teaching not much of a profession. The trouble has been the lack of definite measurements for school products and pupils' abilities. With the advent of educational and mental measurements, a new era in education began. Starting with the earliest attempts about fifty years ago on working out plans to measure different degrees of proficiency in the various school subjects, the testing movement made but little headway then, and no lasting results were achieved. But about thirty years later, Dr. Rice, fresh from his studies in Germany under eminent psychologists, thought of introducing the new idea - that of setting up standards for performance in the common branches of school studies. This idea he carried out by making up a list of fifty words in spelling, tested school after school, and reported that pupils who spent thirty minutes a day in spelling were no better off than those who spent half that long. Severe criticisms were thrust from all sides; school men had always considered that spelling was not to be taught for anything except mental discipline or dexterity; and such a thing was unmeasurable. In spite of all opposition, Dr. Rice succeeded in winning such thoughtful educators as Professor Hanus of Harvard and spreading the flame of the movement. Following the trail of Dr. Rice, Professor Thorndike, ten years later, began to experiment with tests and scales to measure school achievement. The first publication of his handwrit-

ing scale in 1910, marks the real beginning of modern scientific measurement in education. Two years later, the Hillegas composition scale was put out and several achievement tests and scales have since then been devised and standardized.

About the same year that Thorndike was erecting his handwriting scale, the French psychologist, Alfred Binet, was experimenting on a new field of measurement. With a keen psychological insight and wonderful ingenuity, he devised tests that were designed to measure, not school achievement, but general mental capacity. To him is due the greatest credit for being the leader in the second stage of the testing movement - in the realm of mental testing. Finally formulated in 1911, the Binet-Simon intelligence scale led to different revisions and extensions by prominent psychological investigators in America (Goddard, Terman, Kuhlmann, and a few others), and paved the way for other mental tests and scales. Then the outbreak of the World War and America's participation in the great conflict created the need for utilizing the "brain power" of the nation. Seeing this great opportunity, American psychologists offered their services to the government in the spring of 1917. Tests based on the Binet-Simon scale but applied to groups instead of individuals, and individual examinations with revisions, extensions, or modifications of the same scale were used. The success of mental testing in the army was followed by the use of intelligence tests not only in the surveys of school systems throughout the country but also in industrial concerns and employment bureaus. Meeting a much stronger opposition than the achievement test movement, mental testing has progressed in a decade be-

yond expectations and played a significant rôle in American public education.

Invented by Rice and placed on a firm scientific foundation by Thorndike, the measurement of school achievement as the first step in the testing process has won out. Headed by Binet and utilized by American psychologists during the Great War, mental testing as the second step in the testing movement has also triumphed. Indications are that a third step is on the verge of reality and materialization - the stage of ethical, volitional, and emotional measurement.

The progress of the testing movement has been transforming the whole field of education. Grading, promotion, retardation, elimination, methods of teaching, the curriculum, teachers' training, and supervision have all been affected in some way or other. Signs of general school re-organization are in sight. School achievement and pupils' native capacities are now being measured with a good deal of accuracy. Education, now based on scientific investigation, has reached a new age.

(b) Tests and Measurements in American Public Education

Our study has shown both the general and the specific practical uses of tests and measurements in American education. It started with the general value of standards and scales of measurement in any form of organization, whether educational or otherwise. It has tried to show that education is, like material production, a shaping process, one in which the mere factor of growth does not suffice to produce a superior output. It has tried further to establish a relation between the demand for such a superior product

and the need for standards and scales of measurement.

The study has gone further to describe the practical uses to which tests have been put by the American public school administrator. It has illustrated the value of testing by giving the results achieved in several school systems throughout the country by the use of standardized tests. Confined to a treatment of only the most important problems in American public education, the study has tried to demonstrate what an invaluable aid tests are to the school administrator in the solution of such educational problems as grading and promotion, retardation and elimination, methods of instruction and the curriculum, educational and vocational guidance, supervision, teachers' training, and general school organization.

(c) The Rôle of Standardized Tests in Philippine Public School Administration

The study has attempted to show the close resemblance between the American and the Philippine public school system by describing the actual working of both. The study has also drawn the comparison of school problems in the Philippines with those in America. Modeled after the American school and run on the same basis, the Philippine public school has to meet the same problems and apply the same remedies for their solution. After establishing the value of tests and measurements to American public school administration and then showing the close relation of the Philippine public school with the American, the study has gone on to determine the degree of applicability of tests as arranged and standardized in America to Philippine school conditions.

The three preliminary studies made to ascertain this applicability lead to the following general conclusions:

(1) That the criterion of discriminative capacity, so important in determining the satisfactoriness of any achievement or intelligence tests, remains, as shown by the range of scores made both by such a heterogeneous group as the Minnesota students in the Delta 2 examination and by a fairly homogeneous group in the Army Alpha test on the Ohio students;

(2) That the influence of education is considerably less than that of native capacity in determining success in the examination, as indicated by the difference between the scores of the city and provincial teachers of the same educational attainments;

(3) That English, a language practically foreign to the Filipino, does not play so important a part as it would seem at first thought, as shown by the relatively low scores made by the Minnesota and Ohio groups on tests requiring the least amount of use of language;

(4) That tests, as they now stand, differentiate Filipinos into the usual groups of a normal curve distribution and in this respect serve the very purpose that the American administrator has found in the tests;

(5) That finally, changes, if ever made, are few and slight, and must necessarily be, not along lines affecting the use of English, but along the informational, subject-matter, and other phases of the tests; and that additional investigations have to be made to substantiate or verify the conclusions drawn from these preliminary studies, to discover the exact nature of the few

changes needed and work out scientifically reliable and valid substitutes of the same difficulty to the Filipino child as the original items are to the American - changes that will do justice to the Filipino child without reducing the usefulness of the original scale.

To sum up, tests and measurements in their present form, if applied to the Philippine public school system, will play a double rôle: first, they will be of immediate aid to the school administrator in the solution of problems common to both American and Philippine public education; and secondly, they will open a new and wide field of educational research with a view to finally adapting them to Philippine needs and conditions. Whether the immediate aim of offering an invaluable assistance in solving present educational problems in the Islands, or the more remote value of leading to a rich field of scientific investigation in the realm of education, or both are realized - it is obvious in the light of all that has been said in this dissertation that the rôle of standardized tests in Philippine public school administration is of the utmost importance.