

Colorado State Teachers College

BULLETIN

SERIES XX

JUNE, 1920

NUMBER 3

Educational Tests and Measurements

MARVIN F. BEESON, PH. D.



RESEARCH BULLETIN

No. 7

Published Monthly by State Teachers College, Greeley, Colorado
Entered as Second-Class Matter at the Postoffice at Greeley, Colorado,
under the Act of August 24, 1912

STANDARDIZED EDUCATIONAL TESTS
AND MEASUREMENTS AVAILABLE FOR THE
ELEMENTARY SCHOOL SUBJECTS

Marvin F. Beeson, Ph.D.
Professor of Educational Psychology

FOREWORD

The literature and materials of standardized educational tests are increasing so rapidly and becoming so abundant that it is difficult for the teacher or supervisor to find his way through the maze. Hence the purpose of this bulletin is to place in the hands of school superintendents, principals and teachers a brief statement of the nature of the scales, tests and measurements now available for the elementary school subjects, with the available references on each test and a general selected bibliography on the tests for each subject. It was the original intention of the writer to add another division to this bulletin setting forth the purpose and value of educational tests and measurements. Owing to the lack of space that part has been omitted. Most of this has been covered, however, in another bulletin by the writer on "The Value of Standardized Educational Tests to the Teacher," published by Colorado State Teachers College, Greeley, 1920. The general bibliography at the beginning of this bulletin is also added to aid in covering this deficiency.

The tests most widely used and those which seem to be the most promising are indicated by asterisks.

The author wishes to express his indebtedness to Dr. J. D. Heilman and Mr. E. D. Randolph of this institution for helpful suggestions.

The following books on the subject have been consulted:

- Chapman and Rusk: *The Scientific Measurement of Classroom Products*. Silver, Burdett & Co., New York, 1917.
- Monroe, De Voss and Kelly: *Educational Tests and Measurements*. Houghton Mifflin Co., New York and Chicago, 1917.
- Monroe, W. S.: *Measuring the Results of Teaching*. Houghton Mifflin Co., New York and Chicago, 1918.
- National Society for the Study of Education, *Seventeenth Yearbook, Part II—The Measurement of Educational Products*. Public School Publishing Co., Bloomington, Illinois, 1918.
- Starch, D.: *Educational Measurements*. Macmillan Co., 1916.

The most complete bibliographies on the subject are the one by Edna Bryner, included in the *Seventeenth Yearbook of the National Society for the Study of Education*, and the *Library Leaflet No. 2*, April, 1919, Bureau of Education, Washington, D. C.

MARVIN F. BEESON.

Greeley, Colorado, May, 1920.

SELECTED GENERAL BIBLIOGRAPHY ON
STANDARDIZED EDUCATIONAL TESTS AND MEASUREMENTS

- Bachman, F. P.—Subject Matter Standards. *Sch. & Soc.* 10: Oct. 11, 1919, 411-416.
- Bachman, F. P.—Teaching Standards. *Sch. & Soc.* 10: Oct. 18, 1919, 454-457.
- Buckner, C. A.—Educational Diagnosis of Individual Pupils. *Columbia Contrib. to Ed.* No. 98. New York, Columbia University, 1919.
- Beeson, M. F.—The Value of Standardized Educational Tests to the Teacher. *Colorado State Teachers College Bulletin*, Feb., 1920, Greeley.
- Curtis, S. A., and Packer, P. C.—Educational Research. *Jr. Ed. Research* 1: Jan., 1920, 5-19.
- Douglass, Some Uses and Limitations of the Standard Educational Test. *Ed. Admin. & Super.*, Dec., 1919, 475-490.
- Frasier, G. W.—Educational Measurements. *Am. Sch. Bd. Jr.* 60: Jan., 1920, 23-24.
- Gray, W. S.—The Use of Tests in Improving Instruction. *El. Sch. Jr.* 19: 1918, 121-142.
- Hannon, V. A. C.—Improvement in School Subjects Throughout the School Year. *Jr. Ed. Research* 1: Feb., 1920, 81-95.
- Heck, W. H.—Comparative Tests of Home Work and School Work. *Jr. Ed.* Ps. 10: 1919, 153-162.
- Jordan, R. H.—The Use of Tests and Scales as Supervisory Instruments. *Jr. Ed.* 90: Sept., 18, 1919, 255-256.
- Kallom, A. W.—The Importance of Diagnosis in Educational Measurements. *Jr. Ed.* Ps. 10: 1919, 1-12.
- Kruse, P. J.—Overlapping in Certain Grades. *Columbia Contrib. to Ed.* No. 92, New York, Columbia University, 1918, pp. 91.
- Landsittel, A. F. C.—Score Card Method of Teacher Rating. *Ed. Admin. & Super.* 4: 1918, 297-309.
- Lane, H. A.—Standard Tests as an Aid to Supervision. *El. Sch. Jr.* 15: March, 1915, 378-386.
- Monroe, W. S.—Improvement of Instruction Through the Use of Educational Tests. *Jr. Ed. Research* 1: Feb., 1920, 96-102.
- North Carolina, University of—Comparative Results of a Statewide Use of Standard Tests and Measurements, 1918. pp. 24.
- Patri, A.—Educational Measurements. *Sch. & Home (Ethical Culture School, New York City)*, Fall, 1919, 13-17.
- Rogers, Agnes L.—The Scope and Significance of Measurement in Early Elementary Education. *Kindergarten-Primary Magazine* 32: Oct., 1919, 40-44.
- Reeder, J. C.—The Geneseo Scale of Qualities. *El. Sch. Jr.* 20: Dec., 1919, 292-296.
- Smith, E. R.—Scales for the Study of Children's Characteristics. *The Math. Tch.* 12: 1919, 10-16.
- Smith, L. O.—Measuring the Efficiency of Schools. *Nebraska Teacher* 22: Jan. 1920, 218-221.
- Stoops—The Use of Score Cards for Judging Text-Books. *Am. Sch. Bd. Jr.* 1918, 21-22, 56-57.
- Swift, G. C.—Standard Tests for Teachers' Use. *Sch. & Soc.* 8: 1918, 117-118.
- Twiss, G. R.—A Plan for Rating the Teachers in a School System. *Sch. & Soc.* 9: 918, 749-756.
- Van Wagenen, M. J.—Educational Tests and Scales: Their Origin. *Sch. Ed.* 39, Oct., 1919, 44, 46-47.
- Waldron, J.—Standards or Tests by which the Superintendent May Judge of the Efficiency of His Schools. *Catholic Ed. Rev.* 18: March, 1920, 153-163.
- Wilson, G. M.—The Proper Content of a Standard Test. *El. Sch. Jr.* 19: 1918, 375-381.
- Woody—The Teaching of Educational Measurements. *Ed. Admin. & Super.* 5: 1919, 7-14.
- Zerbe, J. L.—Personal Judgments as a Factor in Grading. *Sch. Sci. & Math.* 18: 1918, 405-417.

I—SPELLING

SPELLING, 1897

It is significant that the tests for spelling were the first ones devised. It is one of the most definite of the elementary school branches. The words are usually either entirely right or wholly wrong. The methods of testing the subject readily suggest themselves. The spelling lists and the conditions are easily standardized. The administration of the tests is usually very simple and the results easily scored.

In general there are, with respect to the method of administering the tests, two kinds of spelling tests: those in which the words are dictated in isolated lists, and those in which the words are put in sentences and the sentences dictated. Furthermore, the sentences may be dictated so as to allow the children as much time as they desire for writing them, or they may be timed so that the children will be forced to write at the normal rate of writing for their grade, as determined by standardized tests in writing. This latter method seems to be the more natural, since the pupils are then applying their knowledge of spelling as they usually have occasion to use it, that is, in writing at a normal rate of speed, without the attention concentrated upon the spelling of certain words. The words constituting the spelling test are not known to the pupils, so that attention is not called to the words to be spelled. In fact it is possible to dictate the words without the pupils' knowing at all that it is to be a test in spelling. They might take it to be an exercise in writing or language.

In general more words are misspelled when the words are dictated in sentences than when dictated separately, so that the standards for the same words in the former method must be about 5 to 15% lower than for the latter method. (Courtis and Fordyce; see Monroe, De Voss and Kelly, p. 20.) In the Cleveland Survey, however, no difference was found in the results obtained by the two methods. (See Judd, *Measuring the Results of the Public Schools.*)

Examples of spelling tests and scales which were not intended primarily as timed sentence tests are the Ayres Spelling Scale and the Buckingham Spelling Test. Tests devised as timed sentence tests are the Monroe Timed Sentence Spelling Test and the Courtis Standard Research Tests in Spelling, Forms A and B. The Iowa Dictation and Spelling Tests include both word lists and timed sentence tests.

As to the method of deriving the tests in spelling, three general plans have been followed:

1. Words have been chosen on the basis of their occurrence in several of a large number of standard spelling books. This was the method used in the derivation of the Buckingham Spelling Scale.

2. A study of the words actually used most often in business and personal letters, in newspapers and in standard literature, was made by Ayres for obtaining the thousand words most often used. Jones employed a similar method for the derivation of his material of English spelling, except that he investigated the themes of school children instead of the writings of adults in order to discover the words used by school children. The total number of different words occurring in 75,000 themes written by children of all grades of the elementary school was 4,532.

3. Starch selected for his spelling scale the first non-technical defined word on the even numbered pages of Webster's New International Dictionary, 1910, and divided these words into several lists of 100 words each, according to the length of the words, so that each list is equivalent in difficulty.

RICE SPELLING TEST, 1897

This test is no longer available in printed form and is no longer in general use. It is of great historical importance, however, in that it marks the beginning of the modern movement for scientific measurements in education. Rice made an extensive investigation of spelling by means of his word lists, and as a result he made the assertion that classes devoting fifteen minutes daily to spelling proved to be as good in spelling as many classes devoting a half hour or more to the subject. He found little correlation between the amount of time spent on the subject and the results achieved.

References: Rice, J. M., *The Futility of the Spelling Grind*, *The Forum* 23: March-August, 1897, pp. 163-172 and 409-419.

Tidyman, W. F., *A Critical Study of Rice's Investigation of Spelling Efficiency*, *Ped. Sem.* 22: September, 1915, 391-400.

NATIONAL BUSINESS ABILITY TESTS IN SPELLING, 1912

"The elementary test consists of fifty words chosen from Ayres' list of 542 obtained from the examination of two thousand letters. The advanced spelling test consists of a list of fifty words which are printed incorrectly. In ten minutes the pupil is to write the words correctly."

Address: Sherwin Cody, Business Standards Association, 299 Broadway, New York.

Reference: *Commercial Tests and How to Use Them*, 1919, same address.

*THE BUCKINGHAM SPELLING SCALE, 1913

(University of Illinois)

The first spelling scale still in general use was that devised by Buckingham. His method of procedure was to select words common to two out of five standard spelling books. By this method he secured 5,000 words, which by an elaborate plan he reduced to two lists of twenty-five words each, which were intended to test children of all grades from the third to the eighth. The difficulty of each word has been determined experimentally and weighted so that any number of these words can be given as a test.

The advantage of this test is that the difficulty of each word is known. The greatest disadvantage is that so few words are included in the test, and that these short lists are to be used in all the grades. The test is well standardized, however, and has been used in several of the surveys.

Address: B. R. Buckingham, University of Illinois, Urbana, Illinois.

References: Buckingham, *Spelling Ability: Its Measurement and Distribution*, *Teachers College Contributions to Education*, No. 59, New York, 1913.

Tidyman, W. F., *A Descriptive and Critical Study of Buckingham's Investigation of Spelling Efficiency*; *Ed. Admin. and Sup.* 2: May, 1916, pp. 290-304.

Sackett, L. W., *Measuring a School System by the Buckingham Spelling Scale*, *School and Society* 2: December 11 and 18, 1915, pp. 860-864 and 894-898.

Lewis, E. E., *Testing the Spelling Abilities of Iowa School Children by the Buckingham Spelling Test*. *Elementary School Journal* 16: June, 1916, pp. 556-564.

Nudd, H. W., *The Buckingham Test of the Gary Schools in New York City*, *School and Society* 3: April 8, 1916, pp. 529-532. Also in *Educational Administration and Supervision* 2: May, 1916, 311-316.

THE STARCH SPELLING TEST, 1914

(University of Wisconsin)

This test consists of six lists of 100 words each selected from the dictionary by choosing words at regular intervals. The words are arranged according to length from the shorter to the longer, since it is found that in general the longer words are the more difficult ones. The technical and scientific terms were eliminated, so that the test consists of the non-technical words of the English language.

The advantage of this test is that the score will give one an idea of

the per cent of non-technical words included in his spelling vocabulary of the English language. The disadvantage is that many of them are rarely used, especially by the younger children, and hence are too difficult for a test for most of these. The test does not purport to be a test of words ordinarily used or needed.

The test includes directions for administering and scoring, and standards for the different grades. It has not been used in surveys.

Address: Daniel Starch, University of Wisconsin.

Price: 5c a sheet. One sheet for each examiner or teacher.

References: Starch, Educational Measurements.

Starch, D. The Measurement of Efficiency in Spelling and Overlapping, Etc. Jr. Ed. Ps. 6: March, 1915, 167-186.

Starch, D. Measurement of Ability in Spelling. Jr. Ed. Ps. 88: October 3, 1918, 327-328.

***AYRES SPELLING SCALE, APRIL, 1915**

(Russell Sage Foundation)

This spelling scale consists of a list of 1,000 common words selected by Leonard Ayres from an investigation of personal and business letters and words contained in novels, newspapers and books. The author's original purpose was to select 2,000 of the most common words of the English language, but he found that a few words recurred very frequently and the remainder of the 2,000 words only a very few times, so that he limited his list to 1,000 words. These he called the "foundation spelling vocabulary."

The amount of written material analyzed in securing these words was about 368,000 words. The scale is divided into twenty-six columns, according to the difficulty of the words. The steps of difficulty between all successive columns are equal. The scale is so arranged that columns of varying degrees of difficulty may be used for any grade. Words for spelling tests are usually selected from a column of 73% or 84% difficulty for that particular grade.

The Ayres Scale has been used more widely than any other spelling test, and has proved its value in very many surveys, so that many different comparisons among the results are possible.

Address: The scale may be procured at the price of 5 cents from the Russell Sage Foundation, New York City. (One copy for each examiner or teacher.)

References: It is described in a monograph of the same foundation entitled: A Measuring Scale for Ability in Spelling.

Ballou, F. W., Measuring Boston's Spelling Ability by the Ayres Spelling Scale. School and Society 5: March 3, 1917, pp. 267-270.

Briggs, T. H., and Bamberger, F. E., The Validity of the Ayres Spelling Scale. School and Society 6: November 3, 1917, pp. 538-540.

Scotfield, F. A. Difficulty of Ayres' Spelling Scale as Shown by the Spelling of 560 High School Students. School and Society 4: August 26, 1916, pp. 339-340.

MONROE CLASS RECORD SHEET FOR THE AYRES SPELLING SCALE, 1915

(University of Illinois, Urbana, Illinois)

This is a convenient arrangement for tabulating the scores in the spelling test and finding the median.

Address: Bureau of Educational Measurements and Standards, Kansas State Normal School, Emporia, Kansas.

THE BOSTON SPELLING TESTS, MAY, 1915

(F. W. Ballou, Boston)

This test consists of fifty words for each grade. The words were selected from words most frequently misspelled by the pupils of the seventy elementary school districts of Boston. The measuring of each word is illustrated by a sentence, but the sentences are not to be dictated. Minute directions are in-

cluded for giving and scoring the tests. The test was used in a survey of the Boston schools. The tests are not published for use.

Address: Department of Educational Investigation and Measurement, Boston.

References: Bulletins Nos. 1, 4 and 11, of the Department of Educational Investigation and Measurement, Boston.

JONES, CONCRETE INVESTIGATION OF THE MATERIAL OF ENGLISH SPELLING, DECEMBER, 1915

(University of South Dakota)

This monograph includes a study of the spelling vocabulary of 1,050 pupils from four states. The study was made by examining over 75,000 themes which included about 5,000,000 words. From the study it was found that the total number of different words used was only 4,532. From this it appears that this number of words is the limit of the writing vocabulary of the 1,050 school children from these four states. It is possible, however, that many words whose meanings were known to the children might have been omitted by them in their written compositions because of their hesitancy to attempt to spell the words or to use them in sentences. Very likely many of these would be understood by the children if used by other people. Jones gives the list of words misspelled in each of the grades of the elementary school, and finds that nine-tenths of all the different kinds of misspelled words occurred in the second and third grades.

From the total list he selected 100 "Spelling Demons" with which the children have the greatest difficulty. Four-fifths of these words also occur in the Ayres Scale. The four most frequently misspelled words were found to be: which, their, there and separate. These hundred words may be used as a spelling test, or they may be used as a supplementary list for teaching. The other words included in the investigation, or a selection from these lists, may be used for teaching. From the results of this study Jones has compiled a spelling book called "The Child's Own Spelling Book," which may be obtained from the Capital Supply Company, Pierre, S. D.

It will be observed that Ayres and Jones applied opposite methods in the collection of their material. Ayres' plan was to discover what words were used by business men, and in actual life, while Jones' purpose was to find out what words the children used in their own themes and compositions. Thus the two tests supplement each other, though the former method seems on a whole to be superior, since it inquires what words will be best adapted to the actual needs of life.

Reference: Jones, Concrete Investigation of the Material of English Spelling, Vermillion, S. D., 1914, pp. 27.

THE NEBRASKA SPELLING TEST, 1916

(Dean C. Fordyce, University of Nebraska)

This is a timed sentence test taken from column O of the Ayres Scale. Directions for giving the test, time limits for each grade, and standards are given on the test sheet. The test words are indicated in bold type.

Address: The Nebraska Bureau of Educational Measurements, Lincoln, Nebraska.

***MONROE TIMED SENTENCE SPELLING TEST, 1917**

(University of Illinois, Urbana, Illinois)

There are three of these tests: the first for grades 3 and 4; the second for grades 5 and 6; and the third for grades 7 and 8 and the high school. The tests consist of fifty words each, each chosen from one column of the Ayres list, and arranged in sentences with the time for beginning to dictate each sentence indicated in the margin. The significant words are italicized, and occur in the first or middle parts of the sentences, so that the children may begin a new sentence if they do not have time to complete the one on which they were working.

Complete directions for giving and scoring the tests are given on each test sheet. A class record sheet is included with the test. The test is not

standardized, but tentative standards can be taken from the Ayres Scale, remembering that the children usually spell from 5% to 15% less accurately in the timed sentence tests than in the tests with isolated words. Effective standards are being determined. One copy of the test is needed for each examiner.

Address: W. S. Monroe, University of Illinois, Urbana, Illinois.

***THE COURTIS STANDARD RESEARCH TESTS IN SPELLING, 1917**
(Bureau of Research, Detroit, Michigan)

This test consists of two series of timed sentences, each containing twenty-five words chosen from the Ayres Scale. The first is a preliminary test for training the children in taking dictations; the second part is the actual test. The sentences are dictated to the children according to the time indicated in the margin of the test sheet.

For the second grade lists of words are provided instead of timed sentences on account of the difficulty which second grade children experience with the dictation of sentences.

The method of testing spelling by mistakes in compositions is also discussed. Record blanks and graph sheets and very complete directions for giving and scoring the tests are included with the test material.

Price of envelope for testing forty children, 30c.

Address: S. A. Courtis, 82 Eliot Street, Detroit, Michigan.

THE MINNESOTA SPELLING TEST, 1918

This test, intended for grades 2 to 8, consists of fifty words divided into three lists, two of fifteen words each, and one of twenty words. One list is given each day. The objection to the test is that there are probably too few words, since these are to be given in all the grades beginning with the second. Directions for giving and scoring the tests are included with the material. Class Record Sheet No. 1 contains lines for a tabulation for fifty pupils, and Class Record Sheet No. 2 on the same sheet is for the purpose of tabulating the frequency distribution.

Address: Bureau of Cooperative Research, University of Minnesota.

IOWA DICTATION EXERCISE AND SPELLING TEST, 1918-19

This test consists of three forms, one for each two grades, beginning with the third. Each form is printed on a different color of paper so as to prevent confusion. It is a timed sentence test, and 30 seconds are allowed for writing each sentence. There are ten sentences, each sentence containing two significant words chosen from the Ayres list. The words are underscored in the key provided with the spelling test. The significant words are not known to the children, but they are placed near the middle of the sentence so that the children who write very slowly will have an opportunity to spell them. The children need not be informed that they are to take a spelling test.

In addition to the ten sentences, a list of twenty words is provided for an additional spelling test. Directions for giving and scoring the tests, and a card for class records are also provided.

Address: E. J. Ashbaugh, University of Iowa, Iowa City, Iowa.

***THE IOWA SPELLING SCALE, SEPTEMBER, 1919**
(E. J. Ashbaugh, University of Iowa)

This test really consists of three separate scales, one for grades II, III and IV, one for grades IV, V and VI, and the other for grades VI, VII and VIII. The scale was derived from results of tests in more than one hundred schools in the state of Iowa. The entire series consists of 2,977 words from the written correspondence of Iowa people. "Accuracy of each word was determined on the basis of 200 or more spellings by children in each grade. Thus more than 650,000 spellings were used in each grade. . . . The words were then placed in a separate scale for each grade, the scale being divided into twenty-five steps on the basis of the normal probability curve of dis-

tribution. This means that the difference in difficulty of spelling between the words of any two successive steps is approximately equal. . . ."

"As the increase in spelling accuracy from grade to grade was found to be irregular, it was decided to consolidate the seven grade scales into three separate scales instead of into a single scale as Dr. Ayres did with the 1,000 commonest words." (Reference).

The plan of these scales is very extensive, and they have seemingly been well standardized on Iowa children, so that it promises to become a very useful measuring scale.

One copy of each of the three scales needed for every examiner.

Address: Ernest J. Ashbaugh, University of Iowa, Iowa City.

References: University of Iowa Extension Bulletins Nos. 53, 54 and 55, September and October, 1919.

Foster, The Results of a Recent Spelling Test at the University of Iowa. School and Society 5: April 28, 1917, pp. 506-508.

***BUCKINGHAM'S EXTENSION OF THE AYRES SPELLING SCALE, OCTOBER, 1919**

(University of Illinois)

This extension consists of 505 words which have been added to the Ayres Scale. Most of these words are more difficult than the words Ayres used. Accordingly the scale is enriched at the upper end. These words were also selected in a different manner than Ayres' list. They were chosen according to agreements among spelling books. The words added by Buckingham are printed in italics so that they can be distinguished from the Ayres fundamental vocabulary.

In addition to the supplement to the scale, Buckingham has added valuable comments and explanations as to the method of using the scale, and has shown the difference in the standards to be used when giving the spelling tests in dictated lists and in timed sentences.

Very convenient record blanks in duplicate form and in different colors with explanations for tabulating and for finding the median, are supplied with the scales.

One scale needed for each examiner.

Address: B. R. Buckingham, University of Illinois, Urbana, Illinois.

THE GUHIN "TEST YOURSELF IN SPELLING"

(Date unknown.)

This is a simple spelling test for seventh and eighth grade pupils.

Address: Hub City Supply Co., Aberdeen, S. D.

SELECTED GENERAL BIBLIOGRAPHY ON SPELLING TESTS

- Boston—Department of Educational Investigation and Measurement. Provisional Minimum and Supplementary List of Spelling Words for Pupils in Grades 1-8, Boston, 1914, School Document No. 8.
- Boston—Spelling. Determining the Degree of Difficulty of Spelling Words. Boston, 1915. School Document No. 10.
- Boston—The Teaching of Spelling. Boston, 1916, School Document No. 17.
- Brandenburg, G. C.—Some Possible Secondary Factors in Spelling Ability. School and Society 9: 1918, 632-636.
- Brierley, S. S.—Analysis of the Spelling Process. Jr. Exper. Ped. 4: 1918, 239-254.
- Chadsey, C. E.—The Spelling Problem. Ed. Admin. & Super. 5: 1919, 217-218.
- Cleveland, Ohio—Board of Education. Division of Reference and Research. The Results of a Spelling Test. pp. 19. Bulletin No. 2, 1918.
- Cook and O'Shea—The Child and His Spelling. Indianapolis, 1914, p. 282.
- Courtis, S. A.—Teaching Spelling by Plays and Games. Detroit, 1917, 48 pp.

- Heilman, J. D.—A Study in Spelling. Colorado State Teachers College Bulletin, 1918, Greeley.
- Hill, D. S.—Experimental Studies in Achievement in Spelling. In New Orleans, La., Public Schools. Annual Report 1915-16, Part 2, 35-59.
- Hill, D. S.—Standardized Illustrative Sentences for the Springfield Spelling List. Jr. Ed. Ps. 10: 1919, 285-291.
- Hollingsworth, L. S.—The Psychology of Special Disability in Spelling. Columbia University Contributions to Education, No. 88, New York, 1918.
- Hollingsworth, L. S.—The Psychological Examination of Poor Spellers. Teachers College Record 20: 1919, 126-132.
- Houser, G. D.—The Relation of Spelling Ability to General Intelligence and to Meaning Vocabulary. El. Sch. Jr. 17: June, 1917, 708-718.
- Hunkins, R. V.—An Experiment in Column versus Dictation Spelling. El. Sch. Jr. 19: 689-699, May, 1919.
- Mead, C. D.—The Spelling Ability of Plumas County Pupils. Sacramento, Calif., State Printing Office, 1919, 8 pp.
- Nifenecker, E. A.—Department of Education, Division of Reference and Research. Report on Some Measurements in Spelling in Schools of the Borough of Richmond. City of New York, 1918, 88 pp. Publication No. 16.
- Otis, A. S.—The Reliability of Spelling Scales, involving "Deviation Formula" for Correlation. Sch. & Soc. No. 4, 11, 18, 1916, 676-683, 716-722, 750-756, 793-796.
- Pryor, H. C.—A Suggested Minimal Spelling List. In National Society for the Study of Education, Pt. 1, 16th Yearbook, 1917, 73-84.
- Sackett, L. W.—Disparity of Spelling Scales. Jr. Ed. Ps. 6; Dec. 1915, 627-630.
- Sears, J. B.—Spelling Efficiency in the Oakland Schools. Sch. & Soc. 2: Oct. 9 and 16, 531-537 and 569-574. Full report in Bulletin of Public Schools, Oakland, Calif.
- Starch, D.—The Measurement of Efficiency in Spelling and the Overlapping of Grades in Combined Measurements of Reading, Writing and Spelling. Jr. Ed. Ps. 6: March 1915, 167-86.
- Stuebaker, J. W.—Results of an Investigation of Pupils' Ability to Spell. New York, Newson and Co., 1916, pp. 57.
- Thorndike, E. L.—Means of Measuring School Achievement in Spelling. Ed. Admin. & Super. 1: May, 1915, 306-12.
- Tidyman, F. W.—The Teaching of Spelling, World Book Co., Yonkers, N. Y., 1918.
- Woody, C.—Application of Scientific Method in Evaluating the Subject Matter of Spellers. Jr. Ed. Research 1: 119-128, Feb., 1920.
- Zirbes, L.—An Experimental Evaluation of Method in Spelling. El. Sch. Jr. 19: 1918, 778-798.

II—ARITHMETIC

ARITHMETIC, 1908

The tests in arithmetic are of two kinds, the reasoning tests and the tests in the fundamental operations. The former of course generally include the latter, but the two do not require the same abilities, so that a pupil may do well in a test in arithmetical reasoning and yet fall below the average of the class in speed or in accuracy in the fundamental processes. On the other hand a high performance in the fundamental operations does not necessarily insure superiority in arithmetical reasoning.

It is rather strange that the first standardized tests in arithmetic should have been tests in arithmetical reasoning rather than in the fundamental operations. The latter phase of the subject is simpler, more definite, and certainly easier to test and to score. In fact it is still thought by many authorities that there is not yet a suitable standard test for reasoning ability in arithmetic, while there are several good ones for the fundamental operations.

Stone, the author of the first standardized test in arithmetic, showed in the references cited below that the subject of arithmetic involves a number of specific, and more or less distinct, abilities. Curtis has analyzed further the types of abilities involved in operations in the fundamental processes, and finds at least seven in addition, three in subtraction, six in multiplication and nine in division.

In the following discussion the tests in reasoning ability, the fundamental operations, and the practice tests will be considered successively.

1. Reasoning Tests in Arithmetic

***THE STONE REASONING TEST, 1908 AND 1917** (Teachers College, Columbia University)

There are twelve problems included in this test. The first five receive a value of one point each, the sixth a value of 1.4 points, the seventh a value of 1.2, the eighth 1.6, and the last four a value of 2 points each, if worked correctly. The test is scored by adding the values of the problems worked correctly. Only the method of working the problem is considered, and credit is given for examples partially correct and for those incomplete. Fifteen minutes are allowed for the test.

The test has been widely used in surveys, and so good standards and means of comparison are available for it. One of the disadvantages of it is that the method of scoring is somewhat subjective and open to criticism, since it is difficult to decide how much credit shall be allowed for problems incomplete or partially correct.

Address: Teachers' College, Columbia University, New York City.

References: Stone, *Arithmetical Abilities and Some Factors Determining Them*, Teachers' College Contributions to Education, No. 19, 1908. Stone, C. W., *Standardized Reasoning Tests in Arithmetic and How to Utilize Them*. Columbia University, New York, 1916.

THE BONSER ARITHMETIC REASONING TEST, 1910

This test consists of ten problems divided into two lists of equal difficulty. It is not published for distribution.

Address: Bureau of Publications, Teachers' College, Columbia University.

Reference: Bonser, *The Reasoning Ability of Children of the Fourth, Fifth and Sixth School Grades*, Teachers' College Contribution to Education, 1910.

THE COURTIS STANDARD TEST IN ARITHMETIC, SERIES A, 1911 (Bureau of Research, Detroit)

This series includes two reasoning tests which are described below in connection with the tests in the fundamental operations.

THE BUCKINGHAM REASONING TEST, JANUARY, 1916 (University of Illinois)

This scale appeared as an integral part of the Survey of the Gary and Prevocational Schools of New York City. There are two scales of approximately equal difficulty, consisting of ten problems each. Each question has been carefully evaluated, and the scores are indicated in the margin. Ten minutes are allowed for working each question. The test is not published separately.

Address: Seventeenth Annual Report of the Superintendent of Schools, 1914-15.

Reference: Buckingham, *Notes on the Derivation of Scales of the School Subject with Special Application to Arithmetic*, Fifteenth Year Book of the National Society for the Study of Education, Part I, 1916.

***THE STARCH ARITHMETICAL SCALE A, APRIL, 1916** (University of Wisconsin)

This scale consists of twelve problems of increasing difficulty. The actual values of the problems have been determined experimentally, and the value

of each problem is indicated on the sheet. The approximate value, which is also given, is used in scoring the results. Some of the problems of the Stone Test are also used here. The principle of the test differs from Stone's in that here only the correct answer, not the method of work, is considered in the scoring.

The score of each pupil is the most difficult problem worked correctly by him. If he fails on an example but works a more difficult one, he receives credit for every step passed beyond the point where he first failed, and is penalized for every step on which he failed.

This reasoning test has been used in several of the surveys, so that we have rather good standards for comparison. Directions for administering and scoring the test accompany the material. Pupils are allowed as much time as they need for the test.

Address: Dr. Daniel Starch, University of Wisconsin, Madison, Wisconsin.

Reference: Starch, Educational Measurements, MacMillan & Company.

***MONROE STANDARDIZED REASONING TESTS IN ARITHMETIC, 1918**
(University of Illinois)

This promises to be one of the best reasoning tests in arithmetic. The series consists of three tests; test 1 intended for grades 4 and 5, test 2 for grades 6 and 7, and test 3 for grade 8. There are fifteen problems in each test. The weighted values of the problems have not been statistically determined as yet, but tentative values are given at the present time, with the intention of revising these weights on the bases of the returns from the tests.

The test differs from the previous ones in that each problem is evaluated separately for correct principle and correct answer. The scores are determined by adding the values for the principle and those for the correct answer separately. A score for rate may be obtained by requiring each pupil to draw a line around the number of the problem he is working on at the end of ten minutes, and then continuing the test. Twenty-five minutes are allowed for the test. The test is published in three forms of approximately equal difficulty.

Complete directions for giving and scoring the test, as well as a score key and answer sheet and class record sheet, are included with the test.

The test is being standardized on the basis of the returns of those using it, so that tentative standards are already available.

Address: W. S. Monroe, University of Illinois, Urbana, Illinois.

Reference: Monroe, Measuring the Results of Teaching, Ch. VI.

CLAPP STANDARD SCHOOL TESTS FOR UPPER ARITHMETIC
APRIL, 1918

(University of Colorado; now University of Wisconsin)

This test consists of 25 problems for grades 5, 6, 7 and 8. The test is arranged in the order of increasing difficulty. Directions for giving and scoring the test, and an answer sheet are included with the material. Forty minutes are required for the test, and the score is the total number of problems correctly solved.

Address: Frank L. Clapp, University of Wisconsin, Madison, Wisconsin.

***THE BUCKINGHAM SCALE FOR PROBLEMS IN ARITHMETIC, 1919**
(University of Illinois)

This scale consists of three divisions, of ten problems each, the first for grades 3 and 4, the second for grades 5 and 6, and the last for grades 7 and 8. Each question has been evaluated statistically, and its value is indicated in the margin. The score of the pupil is the value of the most difficult problem for which the correct answer is obtained. For every failure on problems in Division I previous to the most difficult problem solved, 0.3 is subtracted from the score, or 0.2 in Divisions II or III.

The test is issued in very convenient form. There is a square below each problem for the figuring and a smaller square for the answer. Very complete directions for administering and scoring the test, as well as an answer

sheet, a class record sheet, and a detailed score sheet, are included with the materials.

There is no time limit for the test. Tentative standards are obtainable from the following address. The scale is available in several equivalent forms. On the whole it seems to be one of the most promising of all the reasoning tests.

Address: B. R. Buckingham, University of Illinois, Urbana, Illinois.

2. Tests in the Fundamental Operations

THOMPSON'S MINIMUM ESSENTIALS IN ARITHMETIC, MAY, 1908

This consists of a series of thirty oral and written tests in the fundamentals of arithmetic, and in denominate numbers. Two of the latter are tests in elementary denominate numbers, and two others in advanced denominate numbers. The tests in the fundamental operations are of a spiral nature. Sets for drill work are also included with the tests, so that they serve the purpose of both practice tests and performance tests. There is a mechanical device for scoring the papers. They seem to have been originally intended chiefly as teaching devices.

Address: Ginn and Company, New York.

COURTIS STANDARD TESTS, ARITHMETIC, SERIES A, 1911 (Bureau of Research, Detroit)

This is the earliest one of the Curtis Standard Research Tests. It consists of a speed test in copying figures, and speed tests in subtraction, multiplication, division, addition and in the fundamentals combined.

There are also two reasoning tests in the series. Test 8 consists of eight problems in arithmetic to be worked out, and Test 6, the speed test in reasoning, consists of 16 problems. These problems are not to be worked out in full, but the operations used in working them are simply to be indicated. For this reason Test 6 is largely a reading test, and so falls short of its aim. When taken with Test 8, however, the results are enlightening.

Directions for giving the test, and record sheets and answer cards accompany the material.

Address: S. A. Curtis, 82 Eliot Street, Detroit, Michigan.

THE NATIONAL BUSINESS ABILITY TESTS, 1912

The addition test consists of eight examples of three columns each with nine figures in each column, time four minutes. The same sheet contains a subtraction test consisting of fifteen examples, time two minutes. A multiplication test consisting of eleven examples, time allowance three minutes, is also included on the same sheet. Another sheet contains the same addition test again and a test in fractions (short cuts) and percentage. There are sixteen examples in the test for fractions, and the time allowance is five minutes. The test in percentage consists of twelve simple problems to be attempted in two minutes.

With the tests are included complete directions for giving and scoring the tests, standards, answer sheets and record blanks.

Address: Sherwin Cody, 299 Broadway, New York City.

Reference: Cody, Commercial Tests and How to Use Them, 1919.

***COURTIS STANDARD RESEARCH TESTS IN ARITHMETIC SERIES B, 1913**

(Bureau of Research, Detroit)

These tests are undoubtedly the best known and most widely used standardized tests in any subject. One of the chief advantages of the Curtis Tests is that they have been so widely used that very reliable standards have been derived for them. Also very many comparisons can be made with results from various school systems. The test is suitable for a general survey of a town, a school, or a class, but the results are not very reliable for the

performance of individual pupils. It is not a diagnostic test, and was not intended to be used for this purpose.

The test consists of work in the four fundamental operations; a page for each one. The examples are arranged so that all are of equal difficulty. There are 24 examples in the addition test and in all the other tests except multiplication, in which there are 25.

Since all of the examples in each operation are of equal difficulty the pupil's score for rate is the number of examples attempted in the time limit, and the score for accuracy is the percentage of those attempted which are correct. The children score their own tests by means of a printed score card which the teacher reads. An individual score sheet is also provided if desired, so that each child can draw a graph showing his own achievement in comparison with the standard and in comparison with the median of the class. A supervisory graph is furnished with the tests for making tabulations for the whole class. Also a class record sheet accompanies the material, as well as detailed instructions for giving and scoring the tests. The test is issued in four forms of equal difficulty, so that it may be repeated by the same pupils using a different form.

Address: S. A. Courtis, 82 Eliot Street, Detroit, Michigan.

References: Courtis, Teachers' Manual of the Courtis Tests. Also Annual Accounting of the Courtis Research Tests—Department of Co-operative Research, Detroit, Michigan.

Buckingham, R. R.—The Courtis Tests in the Schools of New York City. *Jr. Ed. Psych.* 5: April, 1914, 199-214.

Courtis, S. A.—Courtis Tests in Arithmetic: Value to Superintendents and Teachers. *Fifteenth Yearbook of National Society*, Pt. 1, 1916, 91-106.

Courtis—Measurement of Growth and Efficiency in Arithmetic. *El. Sch. Teacher*, October, December, 1909; December, 1910; March, June, 1911.

Baldwin, B. T.—The Application of the Courtis Tests in Arithmetic to College Students. *Sch. & Soc.* 1: April 17, 1915, 569-576.

Cleveland, Board of Education, Division of Reference and Research—Arithmetical Abilities of School Children as Shown by Courtis Tests. Cleveland, 1917.

Monroe, W. S.—A Report of the Use of the Courtis Standard Research Tests in Arithmetic in 24 Cities. *Studies No. 4*, Kansas State Normal School, Emporia.

FASSETT'S STANDARDIZED NUMBER TESTS, 1914

This consists of four series of tests, one for each of the fundamentals and a separate test for each grade. The tests are arranged on cards in such a way that the pupils can draw a line at the side of the cards and place the answers to the right of the line. The time limit for each exercise is printed on the card, and the directions for using the tests are printed on the binder for each set. The answers are printed on the back of the cards. The tests are probably of greater value as practice tests than for survey purposes.

Address: Milton Bradley Company, Springfield, Massachusetts.

*THE CLEVELAND SURVEY ARITHMETIC TESTS, 1915 (University of Chicago)

This is a series of 15 tests, each of 30 seconds to 3 minutes duration. The total time for taking this test is 22 minutes. The test is given preferably on two different days so that the children take only half at a time. The test is spiral in character; that is, examples of each operation recur in a more difficult form several times. The chief advantage of this test is that it is a diagnostic test. There are many examples of each of several types, so that the results show just what type of example the pupils have difficulty with, and suggest to the teacher a modification of her methods to suit the needs of the class. It has the further advantage of being short and simple.

A score card and directions for giving and scoring the tests are included. The test has not been completely standardized, but the results of the Cleveland and the Grand Rapids Surveys and those of other cities may be used for

comparison. In principle it seems to be one of the best tests. Where it is desirable to express the results of the test in one single score, however, and there is no desire for a diagnosis of the ability of the pupils, some other test should be used.

Address: School of Education, University of Chicago, Chicago, Illinois.

References: Judd, *Measuring the Work of the Public Schools*, Survey Committee, Cleveland, Ohio.

Heckart, J. W.—*Cleveland Survey Tests in Arithmetic in the Miami Valley*. *El. Sch. Jr.* 18: Feb. 1918, 447-457.

BOBBITT'S ARITHMETIC TESTS, 1915

(University of Chicago)

These tests appeared as a part of the Survey of the San Antonio Public School System. They are not published separately.

Reference: *A Survey of the San Antonio Public School System*, School Board, San Antonio, Texas, 1915.

*THE WOODY ARITHMETIC SCALES, SERIES A, 1916

(University of Washington)

This series consists of a separate scale for each of the fundamental operations. As many different types of examples as could be found for each operation are included in the scale. The examples have been arranged in order of increasing difficulty as determined statistically. The addition scale consists of 38 problems, each of a different type, the subtraction scale 35 examples, the multiplication scale 39, and the division scale 36. Twenty minutes are allowed for each operation. The test is not strictly a timed test, as twenty minutes are long enough for most of the pupils to complete the work. The score is the number of the example worked by exactly 50% of the class. A second method of scoring the papers is to find the median of the number of examples worked correctly by each pupil.

There is some difference of opinion with regard to the diagnostic value of these scales. Since there is only one type of each example, a pupil's failure to work the single example of a given type would not necessarily indicate his inability to do all examples of that type. For diagnosis of the class as a whole the scale is no doubt more suited.

These tests have been widely used, so that a number of different comparisons may be made. There are also the tentative standards of the author of the test.

Series B

In Series B some of the examples are omitted, so that the scales are abbreviated forms of Series A. The time limit on this series is 10 minutes for each operation.

Address: Teachers' College, Columbia University, New York City.

References: Woody—*Measurement of Some Achievements in Arithmetic*, Teachers College Contributions to Education, No. 80, Columbia University, New York City, 1916.

Woody—*Measurements of Some Achievements in Arithmetic*. *School and Society*, 4: August 19, 1916, 229-303.

Monroe, W. S.—*An Experimental and Analytical Study of Woody's Arithmetic Scales*, Series B. *School and Society*, 6: October 6, 1917, 412-420.

Theisen, W. W., and Fleming—*The Diagnostic Value of the Woody Arithmetic Scales: a Reply*. *Jr. Ed. Ps.* 9: November, December, 1918.

BOSTON RESEARCH TESTS IN ARITHMETIC

Operations with Fractions, 1916

(By A. W. Kallom)

These tests consist of a series of operations in the addition, subtraction, multiplication and division of fractions. There is a series of six tests in the addition of fractions, each test of a different type of examples; a series

of five tests in the subtraction of fractions, four in the multiplication of fractions, and three in the division of fractions. The tests in multiplication and division are combined in one leaflet, and each of the other two tests is put up in a separate booklet. The tests are not printed for distribution.

Address: Department of Educational Investigation and Measurement, Boston.

Reference: Determining the Achievement of Pupils in the Addition of Fractions, Bulletin No. 7, School Document No. 3, Boston, 1916.

***MONROE DIAGNOSTIC TESTS IN ARITHMETIC, 1916-17** (University of Illinois)

These tests are somewhat on the order of the Cleveland Survey Test. The same general plan is used for constructing it, but it is a much more thorough test, and differs greatly from the Cleveland test in detail.

The series consists of four parts, part 1, operations with integers, intended for grades 4 to 8; part 2, operations with integers of a somewhat more difficult form, intended also for grades 4 to 8; part 3, operations with common fractions in addition, subtraction, multiplication and division, to be given in grades 5 to 8; and part 4, an ingeniously devised test in decimal fractions for grades 6 to 8 and the high school. The answers to the examples in decimals are given and the pupils merely insert the decimal point in the proper place. There are five decimal tests, all in division and multiplication.

It will be noticed that the test is spiral in character, proceeding from the simplest fundamental operations in part 1 to the more difficult fundamental operations in part 2, and then to operations with common fractions and decimals. The time required for giving these tests varies from 30 seconds to 4 minutes each. Each part consists of five or six separate tests on different types of examples. Careful directions for giving and scoring the tests and a class record sheet are included with the materials.

This test promises to be one of the best diagnostic tests which we have. The purpose of the test is to diagnose both the class and the individual. It is well suited for this purpose since there are many examples in each type of problem. By this means the teacher can get an accurate knowledge of the needs of her class and of each individual. It is suggested that this test be used only in making a diagnosis of the school, class or individual and not for a general survey of the school or school system. Some more general test, as the Curtis Test or Woody Test, might be used for determining the standing of the school system as a whole, and then Monroe's diagnostic test given to the children who are above or below the standard, in order to discover the cause of their proficiency or inefficiency and their ability on each type of example.

The disadvantage of the test is that it has not yet been standardized, although the author is deriving standards from the returns of those using the tests, and these will be available soon. However, if a general test is also given, the standards are not so important for diagnostic purposes alone.

Address: Walter S. Monroe, University of Illinois, Urbana, Illinois.

References: Monroe—Measuring the Results of Teaching, Chapters 4, 5 and 6.

Monroe—The Ability to Place the Decimal Point in Divisions. *El. Sch. Jr.* 18: December, 1917, 287-293.

Monroe, W. S.—A Series of Diagnostic Tests in Arithmetic. *El. Sch. Jr.* 19: 1918, 585-607.

THE GUHIN NUMBER TESTS, 1917

This series consists of two tests printed on the same sheet. Both are tests in the simplest addition. The standards are given in terms of the time required for each grade to work the examples on one side of the test sheet.

Address: Hub City School Supply Company, Aberdeen, S. D.

THE GUHIN PRACTICAL MEASURING SET

This is a set for measurement of performance in denominate numbers.

Address: Hub City School Supply Company, Aberdeen, S. D.

WOODY-McCALL MIXED FUNDAMENTALS IN ARITHMETIC, 1918

(Teachers' College, Columbia University)

This is similar to the Woody Scale, but it consists of only one set containing examples in all of the fundamental operations arranged in order of increasing difficulty. The time allowance is twenty minutes, which is supposed to be sufficient to allow practically all children to complete the test. The examples are arranged in spiral form, that is, examples in each fundamental operation recur from time to time in a more difficult form. This test is valuable for a brief survey of a large number of children, or it may be used as a general test when there is not sufficient time to give a more complete test. This scale might also be used to discover in what fundamental operation the pupils are weakest, and then a more complete test, at least for that operation, might be given, such as the Curtis Test or the Woody Scale. This test is a recent one, but promises to be of value in a general way. It is published in several forms so that it may be repeated with the same students by using a different form. Tentative standards and test materials may be procured from the Bureau of Publications, Teachers' College, Columbia University, New York City.

Address: Teachers' College, Columbia University, New York.

3. Practice Tests in Arithmetic

THE THOMPSON MINIMUM ESSENTIALS TEST IN ARITHMETIC, 1908

These tests described above are supplied with sheets for oral and written practice so that they can be used for drill as well as for tests. See description above.

Address: Ginn and Company, New York City.

THE FASSETT NUMBER TESTS, 1914

See description above.

*THE STUDEBAKER ECONOMY PRACTICE EXERCISES IN ARITHMETIC, 1916

(Superintendent, Des Moines, Iowa)

This series consists of a number of cards on which are printed a variety of exercises in the fundamental operations with spaces cut out under the exercises so that the pupils can work the examples through the spaces on ordinary paper placed beneath the cards. There are a number of cards for each operation so as to give the children ample opportunity for practice. The cards for multiplication and division are provided with large spaces below the examples so that all of the work may be included.

Set B 3 for smaller rural schools, the set containing 100 cards, costs \$2.50. A Teacher's Manual for the use of the tests, and a Teacher's Daily Record Sheet and other material are included in the outfit.

Address: Scott, Foresman and Company, Chicago.

*COURTIS STANDARD PRACTICE TESTS IN ARITHMETIC, 1916

(Bureau of Research, Detroit)

These tests consist of 32 lessons in different phases of the fundamental operations, and several tests, printed on cards. Before giving the practice series, lesson number 13 (Test A) is given to the pupils. This test consists of several examples in each of the fundamental operations, and is given in order to find out the achievement of the pupils in each operation. After this the kind of practice tests are given in which the pupils need most help.

Each lesson is printed on a separate card, and the examples are so varied that every type of example in the fundamental operations is included for practice. Each pupil is supplied with a tablet of tissue paper so that the pupil places one of the sheets over a card and merely writes the answers in the proper place. The time required for working one of the lessons is taken, and the lesson is repeated until the pupil is up to standard in time and accuracy. Then the succeeding lesson is prepared in the same way. One sheet of exercises may be used by several pupils.

This is an excellent method of providing drill for the pupils who are below the standard in the fundamental operations. It may also be used to good advantage for the purpose of giving individual instruction either to pupils below standard or below the median of the class or to those above standard in some phases of arithmetic and either behind or up to standard in other elements. By means of this device a teacher can devote her time to the median part of the class, the 50% who can profit most by ordinary class instruction, and give individual instruction in the phases of mathematics needed by those who are behind standard and to those who are above standard. To the latter she might assign a greater number of reasoning problems. In this way each pupil can work at his own rate of speed and the instruction needed by each individual is provided without waste of time, either by the teacher or by the pupil.

Until there are differentiated courses of study, or until some provision is made for individual instruction, this is one of the best plans for giving attention to the highest and the lowest qualities of the class. According to the present method many teachers spend the greater part of their time in instructing the duller pupils in the class, while the other approximate three-fourths of the class suffer as a consequence. It should also be borne in mind that the time devoted to these duller pupils is the time most wastefully spent by the teachers. It requires very much more time and energy on her part to teach these, and in the end they accomplish far less, even with the greater expenditure of time, than the gifted or average members of the class. It is the more gifted after all who are most likely to profit by the instruction they get out of school and make the most use of it.

Address: World Book Company, Yonkers-on-Hudson, New York.

SELECTED GENERAL BIBLIOGRAPHY ON ARITHMETIC TESTS

- Ashbaugh, E. J.—The Arithmetical Skill of Iowa School Children. University of Iowa Extension Bulletin, 1916.
- Ballard, P. B. Norms of Performance in Simple Oral Arithmetic. *Jr. Exper. Ped.* 4: 1918, 236-238.
- Ballou, F. W.—Educational Standards and Educational Measurements. School Document No. 10, 1914. Boston Public Schools.
- Boston. Department of Educational Investigations and Measurements. Provisional Minimal Standards in Addition, Subtraction, Multiplication and Division for Pupils in Grades 4 to 8, 1914, Doc. No. 9.
- Arithmetic. School Document No. 5, 1918.
- Buckingham, B. R.—Notes on the Derivation of Scales in School Subjects: with Special Application to Arithmetic. Fifteenth Yearbook, National Society, Pt. I, 1916, 23-40.
- Counts, G. S.—Arithmetic Tests and Studies in the Psychology of Arithmetic. Chicago, Ill., University of Chicago Press, 1917.
- Courtis, S. A.—Annual Accountings of the Courtis Standard Research Tests. Detroit, Michigan.
- Evans, J. E., and Knoche, F. E.—The Effects of Special Drill in Arithmetic as Measured by the Woody and the Courtis Arithmetic Tests. *Jr. Ed. Ps.* 10: May and June, 1919, 263-276.
- Finley, G. W.—Comparative Study of Three Diagnostic Arithmetic Tests, Research Bulletin, 1920, Colorado State Teachers College, Greeley.
- Haggerty, M. E.—Arithmetic: a Co-operative Study in Educational Measurements. Bloomington, Ind. 1915, 385-508. (Indiana University Studies, 27.)
- Hanus, P. H., and Gaylord, H. D.—Courtis Arithmetic Test Applied to Employes in Business Houses. *Ed. Admin. & Super.* 3: November, 1917, 505-520.
- Jessup, W. A., and Coffman, L. D.—The Supervision of Arithmetic, New York, 1915, pp. 225.
- Kallom, A. W.—The Importance of Diagnosis in Educational Measurements. *Jr. Ed. Ps.* 10: 1919, 1-12.

- Mead, C. D., and Johnson, C. W.—Testing Practice Material in the Fundamentals of Arithmetic. Jr. Ed. Ps. 9: 1918, 287-297.
- Minnick, J. A.—Mathematical Tests—Their Relation to the Mathematics Teacher. Math. Teacher 11: June, 1919, 199-205.
- Moore, R. C.—The Psychology of Number. Jr. Exper. Ped. 4: 1918, 221-236.
- Phillips, F. M.—Relation of Initial Ability to the Extent of Improvement in Certain Mathematical Traits. Ped. Sem. 26: 330-355, December, 1919.
- Rice, J. M.—Educational Research—Arithmetic. Forum 34: October-June, 1902 (initial study).
- Rogers, Agnes L.—Tests of Mathematical Ability—Their Scope and Significance. Math. Teacher 11: June, 1919, 145-164.
- Starch and Elliott—Reliability of Grading Work in Mathematics. Sch. Rev. 21: April, 1913, 254-259.
- Thorndike, E. L.—Measurements of Ability to Solve Arithmetical Problems. Ped. Sem. 21: December, 1914, 495-503.
- Thorndike, E. L.—The Relation Between Speed and Accuracy in Addition. Jr. Ed. Ps. 5: November, 1914, 537 ff.
- Uhl, W. L.—The Use of Standardized Materials in Arithmetic for Diagnosing Pupils' Methods of Work. El. Sch. Jr. 18: November, 1917, 215 ff.
- University of Pittsburgh—Extension Division. A Study of Arithmetic in Western Pennsylvania. Bulletin No. 20, July, 1919.
- Wise, C. T.—A Survey of Arithmetical Problems Arising in Various Occupations. El. Sch. Jr. 20: 1919, 118-136.

III—HANDWRITING

HANDWRITING, MARCH, 1910

Handwriting presents several peculiar difficulties for the examiner. It is not so definite a subject as spelling or arithmetic. In these subjects the elements used are either right or wrong, whereas there are all kinds of handwriting, varying from barely legible to very beautiful handwriting. In general two types of scales have been devised for measuring the quality of handwriting.

According to the first type one aims to decide upon the quality of the writing by the general impression of the specimen. It is assumed that in ordinary life the handwriting of a person is not carefully studied or analyzed to discover its faults and virtues, but it is judged to be good or bad merely by the impression it makes. The different scales of this first type, however, differ among themselves as to the basis upon which, or the distinct qualities by which, ones impression of handwriting is determined. Ayres assumes that legibility alone is the determining factor, and that writing is good if it is easily legible. Thorndike, on the other hand, assumes that we judge handwriting on the basis of legibility, beauty and general merit.

The second type of writing scale attempts to analyze the factors of handwriting into their most important elements and to judge these elements separately. These are the diagnostic and analytic scales, of which there are now several.

In all of the tests for handwriting, speed and quality are scored separately. The score for speed can easily be secured by having the children write a familiar selection from two to four minutes. The letters written in that time by each pupil are then counted, and the average number of letters written per minute is taken as the score for rate of writing. Sometimes the procedure of giving the rate test is modified so that the pupils need not know that they are being timed or are taking a test.

*THORNDIKE'S SCALE FOR MEASURING THE HANDWRITING OF CHILDREN IN GRADES FIVE TO EIGHT, 1910

(Teachers' College, Columbia University)

This was one of the first scales of any kind to be used in the measurement of school subjects. The scale is constructed and the quality of the

specimens to be judged is rated on the basis of "legibility, beauty and general merit" of the handwriting. The specimens were judged and ranked by about 23 to 55 expert judges of handwriting on the basis of the three qualities mentioned above, their judgment being given on one thousand samples obtained from school children in grades 5 to 8. These specimens were divided into fourteen different classes according to quality in such a way that each step was equal in merit to the step between any other two successive specimens of the scale. The unit of the scale equals approximately one-tenth of the difference between the best and the worst of the formal writings of one thousand children. The scale begins with specimen 4, which is barely legible, and extends up to specimen 18, which is a perfect copy book model. The specimens are given in no one system of writing, but every kind of slant and style is used in the scale.

Some of the qualities contain only one example or one specimen, while other qualities are illustrated by as many as three or four examples. This arrangement is a disadvantage, since that quality represented by four examples or four different styles of handwriting is more likely to be assigned often in judging writing than other qualities of which there is only one specimen. The tendency is very strong to select a degree of merit for the score of a given specimen which contains handwriting similar in form to that of the specimen, notwithstanding the advice of the author to be careful to judge solely on the basis of quality without regard to similarity of form and appearance of the specimen to that on the scale. Experience shows that the scores 9 and 11, both of which are illustrated by three examples, are more often assigned than 10, of which there is only one sample.

To score a specimen the examiner should compare it with those of the lower end of the scale, and then move it toward the upper end until a quality of writing is found which is approximately the same as that of the specimen. The number of that quality of writing is taken as the score. Decimals may be used with any score to indicate degrees of difference not shown on the scale.

While it might seem at first that this method of judging handwriting is not more exact than the old method of grading by percentage, repeated tests prove that usually, even without experience in using the scale, the scoring by means of this scale is far more accurate than when done by the old plan. Practice in the use of the scale gives a very great advantage over the old percentile method of grading handwriting. Aside from the accuracy, the scale is of great value because of the objective standards which it gives and the basis of comparison which it offers for the comparing of the grading of one teacher or of the attainment of one school system, or school, or class with that of another. The standard scores obtained from the writing of thousands of other children of each grade in different parts of the United States enable the teacher to compare the work of her school with that done by other schools throughout the country. Before the invention of such scales there was no means of comparing the work of one school with that of others.

If the teacher desires to determine the amount of improvement attained by the use of the scale over the old plan of percentile grading without a scale, she will find a convenient method of doing so outlined in Thorndike's *Teachers' Estimates of the Quality of Specimens of Handwriting*, *Teachers College Record*, Vol. 15, No. 5, 1914, Columbia University, New York City. This contains five problems which may be worked out and fifty specimens of handwriting to be graded by the percentile method and by the use of the scale. The average judgments of a large number of expert judges of handwriting are given as standards and as a basis of comparison. These specimens also provide an excellent means of gaining practice in the use of the writing scale. Every teacher who intends to do much testing by the use of handwriting scales should have some such practice as is provided for by these practice specimens. The working of these problems will convince anyone of the superiority of the writing scales to the percentile plan of grading samples of writing.

The Thorndike Scale has been used very widely in school surveys, so that we have good standards for both speed and quality of writing, and an

excellent basis of comparison with the achievements in other cities. The score for speed is determined by counting the number of letters written by each pupil during a period of two to four minutes and finding the average number of letters written per minute. The median of these scores represents the class score.

Address: Bureau of Publications, Teachers' College, Columbia University, New York City.

References: Thorndike, E. L.—Handwriting; Reprint from Teachers' College Record, Vol. II, No. 2, March, 1910.

Thorndike, E. L.—Teachers' Estimates of the Quality of Specimens of Handwriting. Teachers' College Record 15: November, 1914, 279-291.

Thorndike, E. L.—Means of Measuring School Achievements in Handwriting. Ed. Adm. & Super. 1: May, 1915, 300-305.

***AYRES MEASURING SCALE FOR HANDWRITING, THREE SLANT EDITION, 1912**

(Russell Sage Foundation)

This scale was devised upon the basis of legibility alone. All of the original samples used in the derivation of the scale were taken from the fifth grade alone. The specimens were arranged so that the words were placed out of their natural context, and then submitted to ten competent judges. They were asked to read the specimens, and the time required for reading each specimen was taken with a stop watch. The average time required for reading each of the specimens was taken and those specimens requiring the longest time for reading were placed at the lower end of the scale. The specimens were then arranged in eight degrees of merit so that each step was of equal value, and numbered by tens from 20 to 90. Each quality on the scale contains specimens of three degrees of slant, vertical, semi-slant and full slant or business writing.

To use this scale the specimen to be scored is placed along the scale and compared with specimens of each quality, and the score of that quality on the scale which is most similar to the specimen is assigned to it. Intermediate figures may also be used.

The test is issued in convenient form. It has the advantage of including all three degrees of slant and an equal number of specimens under each quality. It, too, has been widely used in surveys. There is little agreement as to the comparative merits of the Ayres Scale and the Thorndike Scale.

Address: Russell Sage Foundation, New York City.

References: Ayres—A Scale for Measuring the Quality of Handwriting of School Children, Bulletin E 113, Russell Sage Foundation, New York City.

Breed, F. S., and Culp, V.—An Application and Critique of the Ayres' Handwriting Scale. School and Society 2: October 30, 1915, 639-647.

Gray, T. C.—The Training of Judgment in the Use of the Ayres' Scale for Handwriting. Jr. Ed. Psych. 6: February, 1915, 85-98.

King, I., and Johnson, H.—The Writing Abilities of the Elementary and Grammar School Pupils of a City System Measured by the Ayres' Scale. Jr. Ed. Ps. 3: November, 1912, 514-520.

***THE FREEMAN CHART FOR DIAGNOSING FAULTS IN HANDWRITING, 1914**

(University of Chicago)

This chart consists of five separate scales all printed on one sheet for each of the following qualities: uniformity of slant, uniformity of alignment, quality of line, letter formation and spacing. There are three degrees of merit in each one of the scales, qualities 1, 3 and 5. The intermediate qualities 2 and 4 may be used if desired. Each one of the scales contains two or more specimens under each quality. The writing is rated for each one of these five qualities separately, and scored 1, 3 or 5. These scores are added for the total score. A letter gauge consisting of lines of different angles marked on tissue paper is used for determining the uniformity of slant. An align-

ment gauge also drawn on tissue paper is used for determining the uniformity of alignment. Some of the words in the specimen illustrating quality of line are magnified in order to reveal the errors more clearly. In the specimens under letter formation the defects are indicated by small arrows pointing toward them. Deficiencies in spacing are pointed out by marks.

This chart is an analytical scale for the purpose of making a careful diagnosis of the faults in handwriting. This method of judging handwriting is very much slower and more difficult than the method of deciding on the merit of a specimen by the general impression. It is important, however, for the teacher to know what specific faults are involved in the inferior handwriting. For this reason it is suggested the students who rank lowest in the general test be graded in a special test by the Freeman Scale in order to discover the nature of the errors.

Knowing what specific faults are made, however, does not explain the cause of these mistakes. To discover their cause will necessitate watching the pupils carefully while they are writing. This can be done for those who take the Freeman test or those who offer the greatest difficulty in order to discover and correct the cause of the deficiency. The teacher should then give these pupils individual attention and point out to them the cause of the difficulty and give them individual drill for correcting the mistakes.

The Freeman Test is too difficult to use in most surveys, but it is an excellent test for diagnostic purposes. It has been standardized by the author for both rate and quality of writing. Price 25 cents. One copy needed for each examiner.

Address: Houghton Mifflin Company, Chicago.

References: Freeman—An Analytical Scale for Judging Handwriting, *Elementary School Journal* 15: April, 1915, pp. 432-441.

Freeman—The Teaching of Handwriting, Houghton Mifflin Co., Boston, 1914.

Freeman, F. N.—Handwriting Tests for Use in School Surveys. *Elem. Sch. Jr.* 16: February, 1916, 291-301.

Whitney, F. L.—The Pedagogy of the Freeman Handwriting Scale. *Am. Sch. Bd. Jr.* 50: June, 1915, 16, 65-66.

**THE MONROE RECORD SHEET FOR HANDWRITING, USING THE
AYRES SCALE, 1914-15**
(University of Illinois)

On one side of the sheet is a record blank for tabulating the results of the writing tests in rate and quality. On the back of the sheet are general instructions for obtaining samples and rating the specimens, finding class medians, and evaluating the results. Standards are also given.

Address: Bureau of Educational Measurements and Standards, Kansas State Normal School, Emporia, Kansas.

**THE GRAY STANDARD SCORE CARD FOR MEASURING HANDWRITING,
JULY, 1915**
(University of Texas)

This score card divides the essential qualities of handwriting into nine different items. These qualities and the perfect score for each quality are given below:

- | | |
|--------------------|---|
| 1. Heaviness | 3 |
| 2. Slant | 5 |
| Uniformity | |
| Mixed | |
| 3. Size | 7 |
| Uniformity | |
| Too large | |
| Too small | |
| 4. Alignment | 8 |

5. Spacing of lines.....	9
Uniformity	
Too close	
Too far apart	
6. Spacing of words.....	11
Uniformity	
Too close	
Too far apart	
7. Spacing of letters.....	18
Uniformity	
Too close	
Too far apart	
8. Neatness	13
Blotches	
Carelessness	
9. Formation of letters..	26
General form.....	8
Smoothness	6
Letters not closed	5
Parts omitted....	5
Parts added.....	2
Total Score....	—

These scores when added give a total score of 100 points as the highest possible score.

Gray has three forms of these score cards, a large one for the teacher or examiner, a medium size card for scoring the results of each pupil and filing, and a small individual report card for each child to take home. The values assigned to the different qualities have been carefully determined by a statistical study of the returns from seventy-five competent judges of handwriting, consisting of teachers and supervisors of writing, elementary school teachers, and students and teachers of education.

This method of judging handwriting is defended on the ground that the score card has long been used successfully in making judgments of wheat and cattle, and is now being used in judging school buildings, rating teachers and students and other such purposes.

Address: C. Truman Gray, University of Texas.

Reference: Gray—A Score Card for the Measurement of Handwriting. Bulletin, University of Texas, 1915, No. 37.

THE JOHNSON AND STONE HANDWRITING SCALE, FEBRUARY, 1916

This is an analytical scale based upon several factors including movement and an analysis of legibility. There are seven factors considered: letter formation, uniformity of slant, uniformity of alignment, spacing, quality of line, size and degree of slant. Each specimen of the scale is accompanied by an explanation of its faults and merits. The scale is not published for use, but is printed in the following reference.

Johnson and Stone—Measuring the Quality of Handwriting, Elementary School Journal 16: February, 1916, pp. 302-315.

THE MINNESOTA HANDWRITING TESTS, 1917

(Date uncertain)

These tests consist of a blank sheet of paper with spaces for the heading (name, grade, etc.), and a stanza printed at the head, the first stanza of "Sing a Song of Six Pence" to be written for three minutes. The specimens are to be measured by means of the Ayres Scale. Complete directions for giving and scoring the tests are given in an accompanying folder.

Address: Bureau of Co-operative Research, University of Minnesota, Minneapolis, Minnesota.

THE BREED AND DOWNS HANDWRITING SCALE, MARCH, 1917

This scale was made chiefly for local use as a result of a survey of handwriting in the schools of Highland Park, Michigan. The specimens were scored by means of the Thorndike Scale and then five-step scales were constructed for each of four school grades. The construction of such a scale for each grade, no doubt, stimulates great interest among the pupils.

References: Breed and Downs—Measuring and Standardizing the Handwriting in a School System. *Elementary School Journal*, Vol. 17, March, 1917.

*AYRES MEASURING SCALE FOR HANDWRITING, GETTYSBURG EDITION, MARCH, 1917

(Russell Sage Foundation)

This scale differs from the author's other scales, particularly the "Three Slant Edition," in that here all the specimens are written in one slant, a business hand; the specimens are written on ruled paper; the subject matter is the same in all the tests, namely the first three sentences of Lincoln's Gettysburg Address; the time limit is two minutes in all tests; and the specimens are reproduced in blue-black or fountain pen ink. In addition to this the standards for speed and rate for each grade from the fifth to the eighth are given on the scale, and instructions for giving the test and scoring the results are included.

The author says of the scale, "The purpose of the changes introduced in the present edition is to increase the reliability of measurements of handwriting through standardizing methods of securing and scoring samples, and through making numerous improvements in the scale itself designed to reduce variability in the results secured through its use."

This scale was used in the surveys of Cleveland and Gary.

Address: Russell Sage Foundation, New York City.

Reference: Breed, F. S.—Comparative Accuracy of the Ayres Handwriting Scale, Gettysburg Edition. *El. Sch. Jr.* 18: February, 1918, 458-463.

THE ZANER HANDWRITING SCALES AND STANDARDS FOR GRAMMAR GRADES, HIGH SCHOOLS, ETC., 1917

This scale was constructed for judging writing of students of the Zaner method. It really consists of several scales constructed in a similar manner. The first one contains samples ranging from quality 30 progressively by ten points to quality 90. Three methods of marking are indicated, the one by the figures referred to, another by letters A, B, C, D and E, and the third by the explanations Very Poor, Poor, Fair, Good and Excellent. The second scale written on the lower half of the page illustrates styles as well as qualities of writing. Under each specimen is a criticism pointing out the defects or the superior qualities of the writing. This scale is graded from 45 to 95, progressively by tens, as well as by the other two methods. The suggestion is made that the examiner compare the writing with both scales, and if desirable assign the intermediate values indicated in the second scale. Below this scale are suggestions defining and analyzing form, movement and speed and indicating methods of judging these qualities.

Below this is a third scale which consists of samples of the alphabet written in capitals and small letters. Only three qualities, 55, 75 and 95, appear on this scale.

The fourth is a scale of three degrees of quality for figures, the three qualities being represented by the scores 60, 75 and 95.

A fifth scale includes three samples of drill work of two kinds, numbered qualities 50, 70 and 90.

Finally three specimens are given to illustrate the average rating of form and movement combined. These specimens are numbered 60, 65 and 95.

Standards are published on the sheet with the scales. Price 25 cents each. Address: Zaner and Blossom Company, Columbus, Ohio.

KANSAS CITY WRITING SCALE, DECEMBER, 1917

This is another example of a writing scale constructed for local use. Address: Bureau of Research and Efficiency, Kansas City, Missouri.

***COURTIS STANDARD RESEARCH TESTS, SERIES W, WRITING, 1917-18**
(Research Bureau, Detroit)

This test consists of two separate tests, the first one being given on two standardized paragraphs, paragraph 1 and paragraph 2, the former containing 164 letters, and the latter varying in length, since the child's name, age, grade, and school and the date constitute the subject matter to be written. The second test is a test for which the subject matter is names and addresses amounting to 279 letters to be arranged in alphabetical order.

The children practice test 1 on the day previous to the test but do not see test 2. At the time of the test a writing blank is furnished to each child, and the paragraphs to be copied are printed on this blank. The pupils must write each of these paragraphs once. The test is timed by writing a table of numerals from 1 to 20 on the board and counting each unit as 15 seconds. The examiner points to the numerals consecutively as 15 seconds elapse, and the pupils are instructed to write on their papers the number to which the examiner is pointing as they finish writing the paragraph. They immediately begin copying paragraph 2, and then write in the proper blank on the paper the time indicated by the examiner when they complete this. Time is called at the end of five minutes if some are still writing.

On the other side of the paper is a blank for test 2, filing names and addresses which are given at the top. After the test has been explained and illustrated one minute is allowed for study and then two minutes for the test.

The specimens are scored by the Ayres Scale, Gettysburg Edition.

A score card is furnished which gives a count of the letters in tests 1 and 2 and a record blank for each pupil. A class record sheet is included for tabulating the results. Folder B and Folder C give complete explanations of the test and instructions for giving and scoring the test for rate and quality. In Folder C a formula is given for converting the values of the Thorndike Scale into those of the Ayres Scale. Standards are given in Folder C.

Test 1 is a "maximum performance" test. "That is the test is given after practice, the children are told not to hurry and are urged to do their best. Their attention is directed specifically to handwriting." Test 2 is a timed "free choice" test. Attention is not directed specifically to handwriting. The material is new, so that attention must be diverted to reading the copy. Children choose for themselves the rate and quality of the writing.

Address: S. A. Courtis, 82 Eliot Street, Detroit, Michigan.

THE HOLMES PENMANSHIP TESTS, 1917-18
(Harvard University)

This series consists of minute directions and standardized procedure for administering the tests, which are to be scored by the Ayres Scale, Three Slant Edition. The instructions to the examiners, as well as the selection to be written and the directions to the pupils, are all standardized. There are a dictation test and two practice tests of 15 seconds each, followed by a one-minute test and a four-minute test.

The tests are to be repeated at a later date. In the meantime the pupils are required to practice 15 minutes a day in penmanship, but without instruction in penmanship. By this means improvement under standard conditions can be measured, and a standard of improvement based upon a time unit of practice can be derived.

Address: Professor Holmes, Harvard University, Cambridge, Mass.

Reference: Educational Survey of the Public Schools of Brookline, Mass.

THE FRAZIER SCALE FOR HANDWRITING, 1918
(State Normal School, Cheney, Washington)

The specimens for the construction of this scale were selected by rankings on the Thorndike Scale and the judgment of a competent penmanship teacher. Several improvements over the Thorndike Scale have been intro-

duced in the arrangement of the scale. This is another example of the construction of a scale for local use.

Address: State Normal School, Cheney, Washington.

THE REAVIS AND AIKEN SCORE CARD FOR HANDWRITING, 1918

This card is described in the reference below.

Reference: Reavis and Aiken—The Use of a Score Card in Measuring Handwriting. *El. Sch. Jr.* 19: 1918, 36-40.

***THE WISE AND STARCH MEASURING SCALE FOR HANDWRITING, 1919** (University of Wisconsin)

This scale is printed on ruled card board. A business style of handwriting has been used in the better qualities of the scale. The specimens are reproduced in ink, resembling as closely as possible fountain pen ink, and great care has been exercised to make the reproductions as accurate as possible. Other special features of the scale are: uniform test material, including the capital letters, has been used; a continuous series of steps from 0 to 20 has been provided; and the values of the samples have been determined with a high degree of accuracy.

At the top of the scale the authors give the following explanation of the preparation of the scale. "From a total of 627 samples of writing, 227 were selected. These 227 samples were ranked in the order of merit or quality by eleven judges. On the basis of these rankings, 88 samples were chosen from the 227. The 88 samples were then ranked by 100 judges on the basis of which the above 21 samples were selected so that the difference between the successive steps would be such that 75%, or as near that as possible, of the judges considered any given sample better than the next lower one. Finally in order to determine the values of the samples in printed form, the 21 samples as they appear above were ranked by 300 persons. In the case of the 100 rankings of the original hand-written samples, the value of each sample was computed from its comparison with the five next lower samples, that is, from 500 comparative judgments. In the case of the 300 rankings of the printed copies of the original samples, the value of each was computed from its comparison with the two or three next lower samples, that is, from over 600 comparative judgments. Hence, the values of the scale were determined from over 1,100 comparative judgments for each sample obtained from 400 rankings."

Standards for speed and quality based on 6,000 pupils are given above the scale. In addition to this a table for expressing the measurements in terms of the school marks A, B, C, D and E, and a table for expressing these marks in percentages are included.

Address: University Co-operative Company, 504 State Street, Madison, Wisconsin.

References: Starch—A Scale for Measuring Handwriting, *School and Society*, 9: January, 1918, 154-158; 184-188.

Starch, D.—A Revision of the Starch Writing Scale. *Sch. & Soc.* 10: 1919, 498-499.

LISTER AND MYERS' NEW YORK PENMANSHIP SCALE, 1918

(Brooklyn Training School for Teachers)

The authors give at the bottom of the scale the following description of the derivation of it:

"This scale represents the average of the judgments of 21 teachers and writers in the muscular movement system of penmanship, and of four psychologists.

"From nine schools of Greater New York, representative of the best, medium and poorest product of the muscular movement system of writing, 3,550 specimens were selected from at least one entire class of each grade from 3B to 8B inclusive. Each specimen represented one trial from dictation.

"The specimens from each grade were classified into four groups, on the basis of general merit, and the corresponding groups were thoroughly mixed.

Then on the basis of chance, three hundred specimens were selected, with practically the same number from each group.

"According to written instructions each of the 25 judges ranked these specimens on the basis of movement, on the basis of spacing, and on the basis of form.

"The specimens ranking on the average best and poorest were selected for the top and bottom of the scale, respectively. Therefrom the exact numerical rank which the six other samples on the scale should have was determined, and the specimens whose average ranks are the same as these determined positions, or nearest them, were selected.

"With no exception, all the samples on the scale are less than 0.1 from the determined position."

The unique feature of this scale is that each specimen is graded separately for form, movement and spacing. There is a separate specimen for each of these three features for every value. The total score of a sample is the average of the values for form, movement and spacing determined separately. Under each specimen is a line of criticism pointing out the defects of the writing. The scale is one of the newest ones and consequently is not yet widely known. It seems to promise good results, as it makes use of several new principles and improvements.

Address: C. C. Lister and G. C. Myers, Brooklyn Training School for Teachers.

Reference: Lister, C. C., and Myers, G. C.—An Analytical Scale of Handwriting. Jr. Ed. Ps. 9: 1918, 417-431.

SELECTED GENERAL REFERENCES ON TESTS FOR HANDWRITING

- Almack, J. C.—The Writing Ability of Teachers. *Sch. & Soc.* 2: 1919, 389-390.
- Ashbaugh, E. J.—Present Attainment in Handwriting of School Children in Iowa. Iowa City, the University, 1916, 24 pp.
- Boston. Department of Educational Investigation and Measurement. Penmanship. Determining the Achievement of Elementary School Graduates in Handwriting. Boston, 1916, 43 pp. School Document No. 6, Bulletin No. 9.
- Breed, F. S., and Down, E. F.—Measuring and Standardizing Handwriting in the School System. *El. Sch. Jr.* 17: March, 1917, 470-484.
- Breed, F. S., and Culp, V.—Note on the Relation of Legibility and Form in Handwriting. *Sch. & Soc.* 4: 1916, December 2, pp. 870-872.
- Freeman, F. N.—The Teaching of Handwriting. Boston, Houghton Mifflin Co., 1914, pp. 155.
- Freeman, F. W.—Handwriting. Fourteenth Yearbook, National Society, Pt. I, 1915, 61-77.
- Harris, J. H., and Anderson, F. W.—A Study of Handwriting in the Public Schools of Dubuque, Iowa. Bureau of School Measurements, 1916. 10 pp.
- Johnston, J. H.—A Comparison of the Ayres and Thorndike Handwriting Scale. *North Carolina High School Bulletin*, 7: October, 1916, 170-173.
- King, I., and Newcomb, R. E.—Improvement in Handwriting. *Ed. Admin. and Super.* 2: October, 1916, 493-502.
- Koos, L. V.—Determination of Ultimate Standards of Quality in Handwriting for the Public Schools. *El. Sch. Jr.* 18: February, 1918, 423-446.
- Manuel, H. T.—The Use of an Objective Scale for Grading Handwriting. *El. Sch. Jr.* 15: January, 1915, 269-278.
- Manuel, H. T.—Studies in Handwriting. *Sch. & Soc.* 5: March 17, 1917, 327-330.
- Mead, C. D.—The Effect of Exempting Pupils Proficient in Handwriting. *Jr. Ed. Ps.* 10: 1919, 219-229.
- Measuring the Quality of Handwriting. A Symposium. *El. Sch. Jr.* 16: February, 1916, 302-315.
- Nutt, H. W.—Rhythm in Handwriting. *El. Sch. Jr.* 17: February, 1917, 432-445.

- Pechstein, L. A.—Penmanship Scales—Their Merits and Limitations. *Jr. of the N. Y. State Teachers Ass.* 5: March, 1918. 41-45
- Pintner, R.—A Comparison of Ayres and Thorndike Handwriting Scales. *Jr. Ed. Ps.* 5: November, 1914, 525-536.
- Rusk, R. R.—A Class Experiment in Scoring Handwriting. *Jr. Ed. Ps.* 5: September, 1914, 417-418.
- Sackett, L. W.—Comparable Measures of Handwriting. *Sch. & Soc.* 4: October 21, 1916, 640-645.
- Starch, D.—Methods in Constructing Handwriting Scales. *Sch. and Soc.* 10: 1919, 328-329.
- Witham, E. C.—All the Elements of Handwriting Measured. *Ed. Admin. & Sup.* 1: May, 1915, 313-324.
- Witham, E. C.—A Method of Measuring Handwriting. *Am. Sch. Bd. Jr.* 48: May, 1914, 18-19, 72-73.

COMMERCIAL TESTS, 1912

*NATIONAL BUSINESS ABILITY TESTS, 1912

(By Sherwin Cody)

This is an elaborate series of tests designed to cover almost every phase of commercial work, including arithmetic, grammar and spelling. They were designed especially for commercial schools and commercial departments. Many of the tests are applicable to the elementary school, however, and some of them have been listed and described in the proper divisions of this bulletin.

Address: Sherwin Cody, Room 620, 299 Broadway, New York.

Reference: Cody, S.—Commercial Tests and How to Use Them. World Book Co., Yonkers, N. Y., 1919.

V—LANGUAGE AND COMPOSITION

LANGUAGE AND COMPOSITION, 1912

Fewer satisfactory tests have been devised for language and grammar than for the subjects previously described. This is due chiefly to the difficulty experienced in testing these subjects. The chief causes of this difficulty seem to be found in unanalyzed processes and unformulated values. The subject is complex and indefinite. Language includes very many different elements. It cannot, like efficiency in arithmetic, be divided into qualities of speed and accuracy, or any such definite qualities. It includes very many different items, such as correct usage, knowledge of grammatical terms and rules, punctuation, oral and written composition, and possibly general language ability, as evidenced in the ability to rearrange dissected sentences, or to complete mutilated sentences, and the like. Most of the tests attempted to measure only one or two of these important elements.

The other difficulty which arises in connection with the language tests is due to the lack of agreement as to what are the essential purposes of language study. Is it correct usage of words and forms? Is it a knowledge of formal grammar, the ability to classify parts of speech and moods and tenses, and to analyze sentences? Is it the ability to punctuate correctly? Is it proficiency in written or in oral composition? Or is it the development of general language ability, a "language conscience" or "grammatical instinct?" Or is such general language ability attainable at all by education? These are some of the questions which must be answered before we can devise suitable tests. Some authors of tests and scales have sacrificed definiteness of aim to an attempt to combine most of the above elements in a single scale.

In the following discussion the language tests and scales will be considered first, and then the composition scales.

1. Language Tests

THOMPSON'S MINIMUM ESSENTIALS IN LANGUAGE, 1908

This is a rather extensive test, combining the features of a set of practice tests and a set of research tests. They are useful chiefly as teaching devices.

The tests combine the elements of usage in short sentences, pronunciation, punctuation, sentences and nouns, the common irregular verbs, adjectives, adverbs and pronouns, etc. In all the tests there is always a sheet for preliminary oral drill and a second sheet for the written test. Much of the tests is in the form of mutilated sentences, sentences to be completed, or questions to be answered.

Address: Ginn and Company, New York.

THE NATIONAL BUSINESS ABILITY TESTS OF ELEMENTARY GRAMMAR AND ADVANCED GRAMMAR AND PUNCTUATION

1912

These brief tests are all included on one sheet, and accompanying them are directions for giving and scoring the tests.

An Elementary Grammar Test and an Advanced Grammar Test for correct usage are included. Five minutes are required for the former, and ten minutes for the latter. The response called for is to cross out the incorrect form.

The Punctuation Test is also divided into the Elementary and the Advanced Test. In the former, the capitals are to be marked and commas inserted, and ten minutes are allowed for the work. In the Advanced Test the student must insert apostrophes, commas, colons, and semi-colons where needed.

Address: Sherwin Cody, Room 620, 299 Broadway, New York City.

Reference: Cody—Commercial Tests and How to Use Them.

THE COURTIS ENGLISH TEST, SERIES C, 1914

(Bureau of Research, Detroit)

This test proved to be so complicated that its publication has been discontinued. It contains valuable suggestions, however.

Address: S. A. Courtis, 82 Eliot Street, Detroit, Michigan.

Reference: Courtis, S. A.—Standard Tests in English. *El. Sch. Teacher* 14: April, 1914, 374-392.

*THE STARCH GRAMMATICAL SCALE A, 1914

(University of Wisconsin)

This scale contains questions arranged in order of increasing difficulty from step five to step sixteen, with step thirteen omitted, because of failure to find questions of exactly that degree of difficulty. The pupil is to choose between two forms of expression which are given in each step.

A step is passed if three out of four of the questions in each step are answered correctly. The pupil is penalized for every step on which he fails. "In the case of the steps which have only three sentences, all three must be correct in order to pass the step." Those containing less than three sentences can be counted only if none of the steps below that point have been missed.

A sheet of directions and standard June scores and a key for scoring the results are furnished with the tests. This scale has been used in some of the surveys, and is doubtless a very good test when used in connection with some of Starch's other tests for different phases of grammatical knowledge.

Address: Daniel Starch, University of Wisconsin, Madison, Wisconsin.

Reference: Starch—The Measurement of Achievement in English Grammar, *Journal of Educational Psychology*, December, 1915, pp. 615-626, and Starch—Educational Measurements, Macmillan Co.

*STARCH PUNCTUATION SCALE A, 1914

(University of Wisconsin)

This scale is constructed in exactly the same manner as the one just described. It contains ten steps extending from step 6 to step 16. Step 15 is omitted because of the author's inability to find a step of exactly that degree of difficulty. There are from one to four sentences in each step, and the step is counted as passed when the pupil answers three out of four of

the sentences correctly. A key and a sheet of directions for giving the test are included. Standard June scores are given on the directions sheet.

Address: Daniel Starch, University of Wisconsin, Madison, Wisconsin.

Reference: Starch—Educational Measurements (and above reference).

***STARCH ENGLISH GRAMMAR TESTS 1, 2, 3, 1914**

(University of Wisconsin)

Test No. 1 is a test for parts of speech. A selection is given and the nouns, pronouns, adjectives, verbs, adverbs, prepositions, and interjections are to be indicated by standard abbreviations given on the sheet. Three minutes are allowed for the test. The score for each pupil is the number of words correctly designated.

Test No. 2 is a test of knowledge of cases. A selection is given, as in the former case, and the student is to indicate in three minutes time the nominative, possessive and objective cases of as many words as possible, the first letter of each one of these cases being used as an abbreviation. The score is the number of words marked correctly.

Test No. 3 is a test for tense and mood. In the selection given the pupil is to indicate the tense of each verb by the proper abbreviation written above each example. In scoring the results of this test the tense and mood when indicated correctly count as one point each. The total number of moods and tenses given correctly is the score.

Directions and standard scores for grades seven and eight and all the grades of the high school are included with the tests.

Address and references, same as above.

BOSTON TEST IN ACCURATE COPYING, NOVEMBER, 1914

This test was used in the survey of the Boston schools. There is a selection which is to be copied in ink, fifteen minutes being allowed for the purpose. Instructions for giving and scoring the test, a teacher's summary report sheet, and a report blank for each pupil are included. The test is not published for distribution.

Address: Department of Educational Investigation and Measurement, Boston, Mass.

Reference—Determining a Standard in Accurate Copying, Bulletin No. 6, School Document No. 2, 1916.

STARCH ENGLISH VOCABULARY TESTS, 1916

(See Reading)

***TRABUE LANGUAGE SCALES (COMPLETION TESTS)**

B, C, D, E, J, K, L, M, 1916

(Teachers' College, Columbia University)

Scales B, C, D and E are practically of the same difficulty and may be used for the elementary grades. Scales J and K are intended for adults, and Scales L and M for high school students. Scales B and C have been used most often.

Each scale consists of ten mutilated sentences in which the blanks are to be filled in correctly—one word for each blank. In Scale C the most difficult sentence contains seven blanks to be filled in. The relative value of the sentences has been carefully determined, and they are arranged in the order of their difficulty. Seven minutes are allowed for Scales B to E.

A practice sheet for the children below the fourth grade is included with the test. This sheet gives four sample sentences which are to be correctly filled in before the test proper is given.

These scales have been widely used in educational research. They do not seem to test for knowledge of any very definite element of language or grammar, but it is thought that they do test, in a way, "general language ability." It is probable that there is no such general ability.

Directions for administering and scoring the test are given in a booklet.

Address: Bureau of Publications, Teachers' College, Columbia University, New York City.

References: Trabue—Completion-Test Language Scales, same address.

Trabue—Some Results of a Graded Series of Completion Tests. *School and Society* 1: April 10, 1915, 537-540.

Trabue—Completion Tests for Public School Use. *Fifteenth Yearbook of National Society*. Pt. 1, 1916, 52-59.

TRABUE-KELLY, COMPLETION EXERCISE ALPHA AND BETA, 1917
(Teachers' College, Columbia University)

This is an adaptation of the Trabue Completion Scales for individual testing.

Address: Bureau of Publications, Teachers' College, Columbia University, New York City.

Reference: Kelly—Individual Testing with Completion Test Exercises, *Teachers' College Record*, September, 1917.

THE MINNESOTA TESTS IN ENGLISH GRAMMAR, 1916-17

These tests were devised and originally used by Buckingham for the survey of the Gary and Prevocational Schools of New York City. They were later arranged by Haggerty of the University of Minnesota, and issued in their present form. The tests consist of ten questions on different phases of formal grammar. The responses consist of giving the principal parts of verbs, listing parts of speech, analyzing the sentences, using relative pronouns and other difficult parts of speech in sentences, comparing adverbs, giving the case of italicized words, etc.

Complete directions for giving and scoring the tests are included in a separate folder.

Address: University of Minnesota, Minneapolis, Minnesota.

Reference: Buckingham—A Survey of the Gary and Prevocational Schools of New York City, Board of Education, New York, 1916.

***CLAPP'S STANDARD SCHOOL TEST: CORRECT ENGLISH, MARCH, 1918**
(University of Colorado; now University of Wisconsin)

This test is to be used in grades four to eight inclusive. Twenty-five minutes are allowed for the test, and complete directions for giving and scoring the test are included. The test consists of twenty-eight questions including many different elements of English, such as punctuation, capitalization, correct usage, grammatical form, parts of speech, correct forms of words, and the correction of incorrect forms. The difficulty might be mentioned that in practice the standards for spoken English differ from those for written English.

The questions are weighted and the value of each question is placed in the margin of the score sheet. The total number of values equals 100 points, and the score key explains exactly the method for scoring each question. The score of an individual is the sum of the values of the questions answered correctly. Standards are given on the score sheet both for the low and high sections of grades five, six, seven and eight.

This test is a new one and has been used only in local surveys, but the principle seems to be an excellent one, since the test aims to combine many of the elements of grammar.

Address: University of Colorado, Boulder, Colorado.

GREENE, ORGANIZATION TEST, 1918
(University of Iowa)

This test consists of ten dissected sentences, the words and phrases of which are set off in groups and numbered. The test is given to see how well pupils are able to arrange groups of words into sentences in five minutes time. Three practice sentences are given on the back of the test sheet. Instead of writing the words out in the correct order, the response is made by placing in a space provided for the purpose the numbers of the groups of words in

the proper order. A key is provided for scoring, so that the examiner need only compare the numbers on the pupil's paper with those given in the key. Also the value of each sentence is given on the answer key. The pupil's score is the sum of the values of the sentences arranged correctly.

A class record sheet is provided, and contains complete directions for recording the results both of this test and a composition test scored by the Nassau County Supplement to the Hillegas Composition Scale. A correlation sheet is also provided giving complete instructions for correlating the scores of both tests by the Spearman Rank Method (R), and a table for converting this into the Pearson r.

The organization test has the advantage of being extremely simple to use and of giving very definite results. Standard scores are given on the back of the test sheet.

There is need for an organization test for **sentences** in compositions for high school English, but obviously the preparation of such a test would be attended by great difficulties.

Address: S. A. Courtis, 82 Eliot Street, Detroit, Michigan.

***THE CHARTERS DIAGNOSTIC LANGUAGE AND GRAMMAR TEST**

1918-19

1. Pronouns, 1918.
2. Verbs A, 1919.
3. Verbs B, 1919.
4. Miscellaneous.

1. Pronouns

This test appears in two forms, the simplest one intended for grades 3 to 8, printed on one sheet. The pupils are told to read the sentences numbered 1 to 42 carefully to see if they are right or wrong. If they are right, they are to be copied on the blank below; if they are wrong, the correct form of the sentence is to be copied. Directions for giving the tests are printed on the first page of the sheet, and the pupils and teacher are to read these together aloud. Two illustrative examples are given to show the method of work. The test is to be administered in two periods. The more advanced test is given to the children of the seventh and eighth grades. The test is like the first one except that in addition to writing the correct form of the sentences, the pupil must give his reasons for making every change. A blank space is provided for this at the right of the page. This test is also completed in two periods.

A score key and sheet of directions for giving the tests are provided. Record sheets also accompany the test. The incorrect forms inclosed in this test were selected from errors in oral and written language of pupils. The test considers both use and knowledge of formal grammar.

2. Verbs A

This test is intended for grades three to eight, and is similar to the former one for pronouns. There are 40 sentences in the test. The children are to read these. If the form of each sentence is correct, they are to respond by making a cross on the line below. If the verb of the sentence is wrong, they are to write the correct form of the verb on the line below. Directions for giving and scoring the tests are included.

4. Miscellaneous

There are 40 sentences in this test. All the different parts of speech are included. The pupil must read these sentences, and if a sentence is correct he responds by making a cross on the line below; if any form is incorrect the pupil must write the correct form on the line, as in the other tests. Two illustrative examples are given in the instructions, which the children read aloud with the teacher. In grades VII and VIII reasons must be given for the changes.

Address: Bureau of Educational Research, University of Illinois, Urbana, Illinois.

THE BUCKINGHAM ENGLISH GRAMMAR TESTS

These tests are now in preparation and will soon be ready for distribution.

Address: Bureau of Educational Research, University of Illinois, Urbana, Illinois.

2. Composition Scales

There is considerable difficulty in testing ability in composition. The question arises whether more value should be attached to the correct form of composition—that is, correct English, punctuation, freedom from errors, and other technical matters—or whether the general story value of the composition is to receive the greater emphasis, or whether both elements should be taken into consideration. Some teachers grade the composition entirely on the story value, others according to the number of mistakes made in grammar and in form, and others try to combine these two qualities.

Another difficulty is the indefiniteness of the subject. It includes many different elements, so that, in our present lack of careful analysis of the elements comprising the story value of composition, if a score for story value is desired it is only possible to arrive at this by the general impression.

Another difficulty is that of learning to grade compositions by reference to the scale, the same difficulty experienced in the case of handwriting. A great deal of practice is necessary before one becomes expert in the use of such scales, and the difficulty is very much greater than that of judging handwriting. In grading compositions each specimen must be read very carefully, and also each sample on the scale must be thoroughly familiar in order to judge the value of the pupils' specimens. The technique of using the composition scale is so difficult in fact that at first the teachers may mark the composition with a smaller average deviation without the use of the scale than with it. Even if this were always the case, however, it would still be an advantage to use composition scales under certain circumstances. They give the teacher at least an objective criterion on which to base her judgments of different compositions; and thus a basis for comparison is established. The standards obtained from the use of scales are objective, so that a certain quality on a standard scale has a very definite meaning, and this score will convey approximately the same idea of the quality of the composition to different teachers.

Thus these scales enable us to evaluate the work in composition quantitatively, so that the results so obtained can be published in school surveys for the purpose of comparison. Teachers' marks would not have a meaning definite enough to be used in a school survey, since the standards of teachers in grading vary enormously.

Another advantage in the use of composition scales is that the standards obtained may serve as a definite aim or goal for the teacher to work toward.

It is fortunate that after a little practice in the use of scales the results secured by the teachers in grading compositions are more accurate than without the use of the scales. The more the scales are used the more accurate they become as instruments of measurements.

*THE HILLEGAS SCALE FOR THE MEASUREMENT OF QUALITY IN ENGLISH COMPOSITION, 1912

(Commissioner of Education, Vermont)

This scale consists of ten compositions, the values of which have been determined. The values extend from 0-937. The steps are nearly equidistant apart. Three of the ten compositions are artificial productions and the rest were written by high school pupils and college freshmen. The subject is not the same in all the compositions.

To use this scale we simply compare the compositions written by the pupils to whom the test is given with those on the scale and assign the value of the sample which it resembles most closely. Intermediate values may also be used.

This is the earliest and best known of the composition scales. Standards and several good means of comparison for this scale are available since it has been used in many of the surveys. The disadvantages of it are that there is only one sample illustrating each quality; the topic is not the same for each composition, and the values attached are rather cumbersome.

Address: Bureau of Publication, Teachers' College, Columbia University, New York.

References: Hillegas Scale for the Measurement of Quality in English Compositions by Young People. (State Commissioner of Education for Vermont.) Teachers' College Record, 13; September, 1912, pp. 331 to 384.

Kayfetz—A Critical Study of the Hillegas Composition Scale. Ped. Sem. 21: December, 1914, 559-577.

Thorndike—Notes on the Significance and Use of the Hillegas Scale for Measuring the Quality of English Composition. Eng. Jr. 2: November, 1913, 551-561.

Gunther—My Experience with the Hillegas Scale. English Journal 8: November, 1919, 535-542.

Gordon, K.—A Class Experiment with the Hillegas Scale. Jr. Ed. Ps. 9: 1918, 511-514.

***HARVARD-NEWTON SCALES FOR THE MEASUREMENT OF ENGLISH COMPOSITION, SEPTEMBER, 1914**

This series consists of four separate scales, one for each form of discourse: exposition, argumentation, description and narration. All the compositions were written by eighth grade pupils and each scale consists of six compositions with the value of each composition giving both its merits and defects.

The subjects are not assigned, so that this fact makes the composition a little more difficult to evaluate them if all were written on the same subject. There is some advantage in having a separate scale for each form of discourse. Also the plan of pointing out the merits and defects of each sample and of making comparisons between the samples is a decided advantage.

Address: F. W. Ballou, Harvard University, Cambridge, Mass.

References: The Harvard-Newton Bulletin No. II, September, 1914. Scales for the measurement of English composition, Harvard University, Cambridge, Mass.

Brownell—A Test of the Ballou Scale of English Composition. School and Society 4: December 16, 1916, 938-942.

Kayfetz—A Critical Study of the Harvard-Newton Composition Scales. Ped. Sem. 23: 325-347, September, 1916.

***THORNDIKE PRELIMINARY EXTENSION OF THE HILLEGAS SCALE FOR THE MEASUREMENT OF QUALITY IN ENGLISH COMPOSITION BY YOUNG PEOPLE, 1915**

(Teachers' College, Columbia University)

This scale differs from the original Hillegas Scale in several respects. In the first place there are 15 different qualities of composition distinguished. The qualities in the middle of the scale are all illustrated by from 2 to 6 different compositions. The qualities are numbered on the scale of 0 to 95 so that these values are smaller and simpler to deal with than those on the Hillegas Scale. Several of the samples from the original scale are also used.

The advantage of the Thorndike Extension is that several samples are given for some of the qualities and that a greater number of the different qualities are distinguished. On the other hand it is more difficult to learn to use this rather complicated scale.

Address: Bureau of Publication, Teachers' College, Columbia University, New York City.

References: Thorndike—A Scale for Measuring the Merit of English writing. Science 33: June 16, 1911, 935-938.

Thorndike—English Composition: 150 Specimens Arranged for Use in Psychological and Educational Experiments, New York City.

Thorndike—A Scale for Merit in English Writing by Young People. Jr. Ed. Ps. 2: 1911, 361-368.

Johnson, F. W.—The Hillegas-Thorndike Scale for the Measurement of Quality in English Composition. Sch. Rev. 21: January, 1913, 39-49.

Walls, W. A.—The Measurement of English Composition by the Thorndike-Hillegas Scale. Amer. Sch. Bd. Jr. 51: 17, 84, July, 1915.

MINNESOTA DIRECTIONS FOR GIVING AND SCORING OF TEST OF ENGLISH COMPOSITION

(Date not given)

In this test the Hillegas Scale is used for scoring the compositions in grades 2-6. The Harvard-Newton Scale for English Composition is used for grades 7-8. The subject of the composition is "Three Things I Would Do To Improve My School Building and How I Would Do Them."

Ruled sheets are furnished to the pupil for writing the compositions. The subject is printed at the top and the usual blanks—name, grade, etc.—are provided.

The advantage of this test is that a folder of complete direction for giving and scoring the test and for using the scale is given. Another advantage is that the same subject is used for all. The University of Minnesota Record Blank should be used with these tests.

Address: Bureau of Co-operate Research, University of Minnesota, Minneapolis, Minnesota.

THE BREED AND FROSTIC COMPOSITION SCALE, 1917

All the compositions of this scale were written by sixth grade pupils. A part of the story called, "The Picnic," was read to the class and then they were told to complete it in twenty minutes. The advantage of this test is the definiteness of the subject matter and the method of standardizing the conditions of the test, so as to make it as definite as possible. The disadvantage of the test is that it was derived entirely from compositions for the sixth grade pupils. The principle of its derivation seems to be an excellent one. The test is published in the references given below.

Reference: Breed and Frostic—Measuring English Composition in the Sixth Grade. Elementary School Journal 17: January, 1917, pp. 307-325.

*NASSAU COUNTY SUPPLEMENT TO THE HILLEGAS SCALE FOR MEASURING THE QUALITY OF ENGLISH COMPOSITION, 1917

(Prepared by M. R. Trabue, Columbia University)

This scale was devised for the survey of the schools of Nassau County, New York. Ten compositions are included in the scale, one sample for each quality. The first seven compositions were written by children in the elementary schools of Nassau County. The last three were selected from compositions which had been compiled by Thorndike.

The value of each of the compositions, except the last one, is the median rating of 139 judges, using as the basis of their rating the Hillegas Scale. The value of the last composition is that given it in the Thorndike Extensions of the Hillegas Scale. The values of these compositions extend from 0 to 9, decimals being used for the intermediate values. The subject of the first seven compositions is: "What I Should Like To Do Next Saturday." There is no definite subject for the last three.

The advantage of this scale is its simplicity. There are only 10 qualities of compositions and there is only one sample for each quality. The values are also very carefully determined. The use of the same subject in most of the compositions is another advantage. Standard medians may be found in the references below.

References: Trabue—Supplementing the Hillegas Scale, Teachers' College Record 18: January, 1917, pp. 51-84. The Nassau County Scale Survey, Nassau County, New York.

Theisen, W. W.—Improving Teachers' Estimates of Composition Specimens with the Aid of the Trabue Nassau County Scale. *School and Society* 7: 143-150, February 2, 1918.

Witham, E. C.—Notes on the use of a Composition Scale. *Jr. Ed. Ps.* 10: November, 1919, 461-462.

***THE WILLING SCALE FOR MEASURING WRITTEN COMPOSITIONS**
1917

(Class Records Sheet by W. S. Monroe)

This scale consists of eight different compositions. The qualities are numbered by tens from 20 to 90, 20 meaning from 15 to 24.9, for example. The subject is the same for all of the compositions, namely, "An Exciting Experience." Several specific topics are mentioned in order to illustrate to the children just what kind of subject is desired.

Complete directions for using this scale are given in the bottom paragraph. The unique feature of this scale is that the compositions are rated for both story value and form value. The rating for form value is done by carefully marking all errors in grammar, punctuation, capitalization and spelling. These are counted and the total is multiplied by 100. This is divided by the number of words in the composition. The quotient is the index of the form value.

This scale has several advantages; the definiteness of subject assigned, the simplicity of its arrangement, the rating for both story and form value, the fact that the directions for using the scale are included, and the small number of degrees of merit to be distinguished. Standards are also included on the scale. This scale was used in the survey of the Denver Schools and that of the Grand Rapids Schools.

Address: Bureau of Educational Tests and Measurements, Kansas State Normal School, Emporia, Kansas.

References: School Survey of Grand Rapids, Michigan, Board of Education, Grand Rapids, 1916.

Denver School Survey, 1916.

Willing, M. H.—The Measurement of Written Composition in Grades IV to VIII. *Eng. Jr.* 7: 193-202, March, 1918.

THE COURTIS STANDARD RESEARCH TEST IN COMPOSITION, 1917-18
(Bureau of Research, Detroit)

This is a folder containing detailed instructions for giving a composition test which is to be scored for rate by the method indicated by the author, and for quality according to the Trabue Modification of the Hillegas Scale.

Address: S. A. Courtis, 82 Eliot Street, Detroit, Michigan.

***THE MINNESOTA ENGLISH COMPOSITION SCALE, 1920**

The reference describing the derivation of this test has not yet been published, and so no description of the scale is given here.

Address: Bureau of Co-operative Research, University of Minnesota, Minneapolis, Minnesota.

Reference: Van Wagenen, M. J.—The Minnesota English Composition Scale, Its Derivation and Validity. (This article will be published in the *Journal of Educational Research* some time in 1920.)

**SELECTED GENERAL BIBLIOGRAPHY ON TESTS FOR ENGLISH,
LANGUAGE AND COMPOSITION**

Barthelmess, H. M.—Determining the Achievements of Pupils in Letter Writing. *Bulletin No. XVI* of the Department of Educational Investigation and Measurement, Boston, Mass., 1918, pp. 35.

Bonser, F. G., Burch, L. H., and Turner, M. R.—Vocabulary Tests as Measurements of School Efficiency. *Sch. & Soc.* 2: November 13, 1915, 713-718.

Boston. Dept. of Ed. Investigation and Mes. English: School Document No. 6, 1918.

- Brandenburg, G. C.—Psychological Aspects of Language. Jr. Ed. Ps. 9: 1918, 313-332.
- Breed, F. S., and Haggerty, M. E.—The Measurement of Improvement in English Composition. Eng. Jr. 6: October, 1917, 515-527.
- Charters, W. W.—Minimal Essentials in Elementary Language and Grammar. Sixteenth Yearbook, National Society, Pt. 1, 1917, 85-110.
- Charters, W. W.—Constructing a Language and Grammar Scale. Jr. Ed. Research 1: April, 1920, 249-257.
- Chicago University High School, English Dept.—Differentiating Instruction in Ninth-Grade English. Sch. Rev. 27: December, 1919, 748-756.
- Curtis, S. A.—The Value of Measurements. Eng. Jr. 8: 1919, 208-217. Detroit, Mich.—English Composition for Use in the Detroit Public Schools. September, 1917. Board of Education, 158 pp.
- Hosie, J. F.—The Essentials of Composition and Grammar. Sch. & Soc. 1: April 24, 1915, 581-587. Also in Fourteenth Yearbook, Nat. Soc. Part I, 1915.
- Hudelson, E.—Standards and Measurements in English Composition. Second Ind. Conf. on Ed. Mes., 13: 1915, 115-122.
- Hudelson, E.—Some Achievements in the Establishment of a Standard for the Measurement of English Composition in the Bloomington, Indiana, School. Eng. Jr. 5: November, 1916, 590-597.
- Johnson, R. Q.—The Persistency of Error in English Composition. Sch. Rev. 25, October, 1917, 555-580.
- Jones, R. G.—Standard Vocabulary. Fourteenth Yearbook National Society, Pt. I, 1915, 37-42.
- Mahoney, J. J.—Standards in English. Yonkers, N. Y., 198 pp., World Book Co., 1917.
- Maxwell, W. H.—Report of Committee on Tests and Standards of Efficiency in Schools and School Systems: English Grammar. Proc. N. E. A., 1915, 565-573.
- Parker, F. E.—The Value of Measurements. Eng. Jr. 8: 1919, 203-208.
- Pintner, R.—The Measurement of Progress in Language Ability. Jr. Ed. Ps. 9: 1918, 270-277.
- Rice, J. M.—The Results of a Test in Language. Forum 35: October, 1903, 269-293 and March, 1904, 440-457. Also in his Scientific Management in Education. New York, 1913.
- Sackett, L. W.—Comparable Measures of Composition. Sch. & Soc. 5: February 24, 1917: 233-239.
- Sears, Isabel, Diebel & Amelia—A Study of the Common Mistakes in Pupils' Oral English. Elem. Sch. Jr. 71; September, 1916, 44-54.
- Stark, W. E.—Measurement of Eighth-Grade Composition. Sch. & Soc. 2, August 7, 1915, 208-216.
- Stoddard, W. E.—A Comparison of the Hillegas and Harvard-Newton Scales in English Composition. Ped. Sem. 23: December, 1916, 498-501.
- Towne, C. F.—Making a Scale for the Measurement of English Composition. El. Sch. Jr. 19: September, 1918, 41-53.
- Vocabulary Tests as Measures of School Efficiency. Sch. & Soc. 2: November 13, 1915, 713-718.
- Ward, C. H.—The Scale Illusion. Eng. Jr. 6: April, 1917, 221-230.
- Wolfe, C. S.—The Topeka Scale for Judging Compositions. Illinois Ass. of Teachers of Eng., Bul. 9, March, 1917, 1-12.

VI—DRAWING

THORNDIKE SCALE FOR THE MERIT OF DRAWINGS BY PUPILS EIGHT TO FIFTEEN YEARS OLD, 1913

(Teachers' College, Columbia University)

This scale is similar to the handwriting scale. The evaluation of the different samples are the results of judgments by 400 artists, "teachers of drawing, and men expert in education in general." The degrees of merit extend from 0 to 17. The drawing to be judged is compared with the scale

and the evaluation attached to the sample which it most nearly resembles is taken as the score.

Address: Bureau of Publications, Teachers' College, Columbia University, New York City.

References: Thorndike—The Measurement of Achievement in Drawing, Teachers' College Record 14: November, 1913, pp. 345-383.

Childs—Measurement of Drawing Ability of 2,177 Children in Indiana City School Systems by the Supplementary Thorndike Scale, Journal of Educational Psychology 6: pp. 391-408, September, 1915.

***THE WHITFORD ART TESTS, 1919**
(University of Chicago)

According to the author, the two fundamental results of art education are ability to recognize and appreciate art quality and ability to draw or describe things graphically. Consequently two types of tests have been included in this series. Test No. 1 deals with ability to appreciate art quality, which is a mental product of art training. Test No. 2 is a test of ability to draw, i. e., "to represent graphically ideas and things of artistic quality."

Test No. 1 consists of a series of 14 problems, each in selection or discrimination between 3 or 4 superior and inferior art considerations. The best drawing of each group is to be marked by the pupils. Only one answer for each problem is correct. The time generally required for this test is less than 15 minutes. The scoring involves no difficulties, since $7\frac{1}{7}$ points are subtracted from 100 for every incorrect response.

Test No. 2—the drawing test—consists of 7 problems. These are divided into four groups. The scale is reproduced below.

I. To test proportion drawn from specification.

1. Draw freehand a small-scale rectangle in the proportion 5 to 12.
2. Draw freehand a small-scale triangle in the proportion 4 by 4 by 2.

II. To test proportion in representing a given figure (copy).

3. Make a small-scale drawing of the map on board (United States).

III. To test drawing of curved lines.

4. Copy the curved lines drawn on the board. (Instructor draws three groups of well selected curves. This enables the pupil to see how they are drawn.)

IV. To test representation (from memory and from objects).

5. Draw from memory a horse (side view).
6. Make a sketch of the chalk box (on desk).
7. Make a sketch of the waste basket (on chair).

This test is scored according to a rating scale of ten different standards of attainment for each problem of the test representing approximately equal steps of difficulty from the poorest drawing to the best found in all the grades.

The test has been standardized on the basis of the results from 1,000 children in three schools. Thirty persons, supervisors of art in public schools and instructors, have aided in the preparation of this scale. A revision of the tests with still simpler technique of scoring based upon results from a larger number of children is being made.

Address: The Prang Company, Chicago.

Reference: Whitford, W. G.—Empirical Study of Pupils in Ability in Public School Art Courses. *El. Sch. Jr.*, September and October, 1919.

GENERAL BIBLIOGRAPHY ON TESTS FOR DRAWING AND ART

Ayer, F. C.—The Psychology of Drawing. Baltimore, Md., 1916, 186 pp.
Eastern Arts Association—Report of Committee on a Scale for Drawing. Proc. 1916, 180-181.

- Manuel, H. T.—Talent in Drawing. Public School Publishing Co., Bloomington, Ill., 1920.
- Pintner, R.—Aesthetic Appreciation of Pictures by Children. *Ped. Sem.* 25: 1918, 216-218.
- Rugg, H. O.—A Scale for Measuring Free-Hand Lettering. *Jr. Ed. Ps.* 6: January, 1915, 25-42.
- Whitford, W. G.—Empirical Study of Pupil Ability in Public School Art Courses, Part I. *El. Sch. Jr.* 20: 1919, 32-46.

VII—READING

READING TESTS, 1914

Reading is a difficult subject to test because there are so many factors involved, and these elements are usually rather indefinite. Some of the elements involved in reading are speed, comprehension, word knowledge, expression, pronunciation, emphasis and readiness to use printed material to advance personal ends, to solve personal difficulties, and to gain information. The purpose of teaching elementary reading is to enable children to get the thought from the printed page as quickly and as accurately as possible, and in the upper grades to develop taste and appreciation for reading, and to aid the children to gain information. Usually in our teaching of reading we lose sight of this aim and stress the secondary aim of reading, namely, to pronounce the words correctly and to learn to read aloud with expression. In reality the most important use we have for reading is not to read aloud with expression but to read silently and get the thought as rapidly and as accurately as possible. Most teachers should stress this factor of thought-getting, in other words, silent reading, much more than they do and lay less stress on oral reading. We have occasion to get the contents from a book much more often than to read it aloud, and for that reason the major emphasis in our school-room instruction should be placed on silent reading for the purpose of getting the thought. Reading might also be given for the purpose of increasing the vocabulary of the child. If this is made a conscious aim of the instruction in reading, it will be found that the pupils do increase their vocabularies very greatly by this method, but if left as a matter of incidental gain it will be found that little profit results from these methods.

It is found from the results from reading tests in many different school systems and from psychological tests in the laboratory that the usual idea that the slow reader gets the context of the page more accurately than the rapid reader is not true to fact. There is a little disagreement on the correlation of these two abilities, speed and comprehension, but many authorities agree that usually the rapid reader also gets the thought of what he reads better than the slow reader. There are several reasons for this, namely: greater concentration of attention when working rapidly, also the general tendency of the intelligent person to work faster as well as more accurately than the duller one, and the general correlation found in all mental tests between the rapidity and accuracy of the mental processes. Practically all authorities, both psychologists and those who have used the tests widely, agree that at least the moderately rapid reader has a better comprehension of the subject than the slow reader. It is possible that the very rapid reader skims over the subject matter without getting the thought, that is without actually reading. Results of psychological tests and educational tests also point out the correlation between speed of reading and retention. The rapid reader usually retains more and for a longer time than the slow reader for similar reasons as those above. One of the chief duties of a teacher in instruction in reading then should be to develop speed and accuracy of silent reading and to spend less time in drill work on the mechanical phases of the subject. That much time is wasted in our ordinary methods of teaching is proved by the fact that in most of our tests and measurements and surveys, it has been observed that pupils make little progress in reading after the third or fourth grade. The time devoted to reading after that point is really worse than wasted because the children are forming

habits of carelessly repeating, in a mechanical way, subject matter already known. Too much stress on oral reading causes the children to pay undue attention to the mechanics of reading so that the tendency may persist to pronounce each word with the lips or in the mind when reading silently. Much of the time given to reading might profitably be devoted to other subjects, or to other kinds of reading which will be found useful. Instead of so much drill work then, the children in the upper grades especially should have a great deal of useful reading, parallel reading, not only of a literary nature, but of many other kinds. Since most of the reading done in actual life is not literary material but magazines, newspapers, scientific books, informational works of all kinds, much of the cursory reading in school should be done in these other fields, not distinctly literary. Such a course would prepare the children better for the actual demands of life.

There are two distinct divisions of reading tests, those for oral reading and tests for silent reading. The tests in silent reading are again divided into several different kinds. They may be tests of word knowledge or vocabulary, or they may be tests of comprehension of the material. Of these latter there are again several different types.

One type is those tests which demand that the speed of reading be scored according to the number of words read per minute or per second and then that comprehension be scored either by reporting orally to the examiner the contents of the passage read or by writing out a report of all that was read. This kind of test may be scored for comprehension either by counting the number of words written after eliminating all repetition of words and redundant words, or by referring to a score key on which are indicated all the ideas contained in the passage. The ideas given in the report are then counted, and the number of ideas given constitute the score for comprehension.

Second, the comprehension of tests for the understanding of sentences may be scored by asking questions on the contents of the passage read and weighting these questions by numbers proportionate to the difficulty. Many of the tests combine these two methods, requiring the pupils to write a report of the passage and also to answer questions bearing on the passage read.

Third, there are certain performance tests, or tests chiefly of the nature of performance, which demand that the child shall read the passage carefully and then make certain marks or symbols or follow certain directions to show that he has understood the contents of the selection and the directions given in the selection. Such tests as these are similar in some respects to the test for following directions in the Army Mental Tests, Scale Alpha. These performance tests may be scored in two ways; either the rate of reading and the understanding may be included in one score, or there may be a separate score for the rate of reading and for comprehension. The Kansas Silent Reading Tests follow the first plan, while the Monroe Silent Reading Tests follow the second plan.

1. Silent Reading Tests

*BROWN SILENT READING TESTS, SEPTEMBER 1, 1914

(State Normal School, Oshkosh, Wisconsin)

This test consists of an interesting reading selection for all the grades from the third to the eighth, the same selection being used for all the grades. There are three forms of equal difficulty provided so that the same tests can be repeated several times by the same pupils. The time required for giving the test is exactly 60 seconds. The pupil is asked to draw a line round the last word read when time is called. The number of words read during the minute are taken as a score for speed. The pupil is then required to write as much as he remembers of what was read. This reproduction is then compared with a key which is furnished with the test and which contains in italics all the principal ideas occurring in the selection. The pupil's paper is scored by the number of ideas correctly reproduced according to the key. The principle of this test has been used in many succeeding tests, since Brown's was the first successful test devised for reading. There are only tentative standards for this test at the present time. The advantage of this method

of scoring is that it gives an objective basis for judging and marking the papers. It is somewhat slow, however. The use of the same selection for all the grades might limit the usefulness of this test somewhat.

Address: H. A. Brown, president of the State Normal School, Oshkosh, Wisconsin.

References: Brown—The Measurement of Ability to Read, New Hampshire Department of Public Instruction, Bulletin No. 1, Second Edition, 1916, Concord, New Hampshire.

Brown, H. A.—The Measurement of the Efficiency of Instruction in Reading. *El. Sch. Tch.* 14: June, 1914, 477-490.

COURTIS RESEARCH TESTS, SERIES C, ENGLISH, 1914

(Bureau of Research, Detroit)

This test was found to be so complicated to use that its publication has been discontinued. The author's reading test described later has taken its place.

Address: S. A. Courtis, 82 Eliot Street, Detroit, Michigan.

*THE GRAY SILENT READING TEST, 1914

(University of Chicago)

This is a very unique test of silent reading. Each child is tested individually. This makes the test both very accurate yet very difficult to give. There are three selections. Grades 2 and 3 are tested on "Tiny Tad," grades 4, 5 and 6 on "The Grasshoppers," grades 7 and 8 on "Ancient Ships." The selections are printed on cards and are so arranged on the cards that the selection is printed in three parts. For "Tiny Tad" the middle portion contains just 100 words. The teacher observes the child read, and by the stop watch takes the exact time when the pupil's eyes shift from the first division to the second division at the top of the card, and then takes the time again when the child's eyes shift from the bottom of the middle division to the top of the last division of the selection. This gives the exact time required for reading 100 words. The number of seconds required for reading the middle column is recorded. This is the score for rate. The middle column of the other two selections contains 200 words, and so the number of seconds required for reading the second column must be divided by two in order to get the score for rate for those two selections.

If the pupil has been reading "Tiny Tad," the teacher asks him to tell the story as well as he can, and writes down the pupil's exact report. When he discontinues speaking he is asked if he remembers anything else. For the other two selections the pupil writes out on the record sheet a report of what he remembers. Also on the back of the sheet are 10 questions on the story read. Each question counts 10%, so that there are two scores for quality of silent reading, the pupil's report of what he read, and the answers to the 10 questions. To score the reproduction, the teacher should check from the pupil's report all wrong statements and all repetitions, and count the remaining words. The percentage that these words are of the total number of words in the selection is the score. For the questions, the score is determined by giving each one of the 10 questions answered correctly a score of 10%. The average of the score for reproduction and the score for answering questions is the final score for quality.

The same record blank is used for both Gray's Oral Reading and Silent Reading Tests. On the back of the score sheet are included explicit directions for giving both tests, and directions for scoring and tabulating the results. Standards for both oral and silent reading are given on the score sheet.

This test has several advantages, the chief being the accurate method of testing reading by testing each individual separately, the ingenious methods of securing the score for rate, and the fact that each pupil is tested for quality of reading by two different methods. The disadvantage is the time required for giving the test. This test was used in the Cleveland Survey, so that there is a good basis of comparison with the results in other schools.

Address: Schools of Education, University of Chicago.

References: Gray—Methods of Testing Reading; Elementary School Journal 16, January, 1916, pp. 231-246, and February, 1916, pp. 281-298.

Gray, W. S.—Reading in the Elementary Schools of Indianapolis. *El. Sch. Jr.*, 1919, 19: 336-353; 419-444; 506-531; 608-627.

Judd—Measuring the Results of the Public Schools, Cleveland Survey.

Gray, W. S.—Studies of Elementary School Reading Through Standardized Tests. Chicago, 1917.

Gray, W. S.—A Co-operative Study in Reading in Eleven Cities in Northern Illinois. *El. Sch. Jr.* 17: 250-265, 1916.

White, C. W.—A Study in Reading in Indiana Cities. Fourth Conference on Educational Measurements. Indiana University, 1917.

*THE KANSAS SILENT READING TESTS, 1914

(Devised by F. J. Kelley, State Normal School, Emporia, Kansas)

This test has been one of the most widely used tests for reading ability. There are three different tests in the series; two for the elementary school and one for the high school. Test No. 1 is to be given to grades 3, 4 and 5; test No. 2 to grades 6, 7 and 8; and test No. 3 to the high school pupils. Exactly five minutes are allowed for giving the test. On the outside page are blanks for the heading, "pupil's name, date, age, grade, etc.," directions to be read in concert by the pupils and teacher, and an example illustrating the tests. After the pupils understand what they are to do, a signal is given to turn over the page and begin with test No. 1. They do as many as they can in the time allowed. The test consists of 16 paragraphs which are to be read and responded to by an answer, or a direction followed. The tests are weighted, and the evaluation of each paragraph is given in the left margin. The score of each pupil is the sum of the values of the paragraphs answered correctly. An answer sheet is furnished with the test, and each question must be answered absolutely correctly in order to receive credit. There is no separate score for rate, but the score for comprehension is intended to include rate of work also.

Instructions for giving and scoring the tests and for finding the median, class record blanks, and an answer sheet are included with the tests. We have very complete standards for this test. It has been widely used in surveys so that there is an excellent basis of comparison with other school systems. The standard medians given on the score sheet were obtained from over 100,000 children.

Address: State Normal, Emporia, Kansas.

References: Kelley—The Kansas Silent Reading Test, Bureau of Tests and Measurements, Publication, New Series 4: 1915, pp. 37.

Kelley—The Kansas Silent Reading Tests, *Journal of Educational Psychology*, Volume 7, February, 1916, pp. 63-80.

Monroe—A Report on the Use of the Kansas Silent Reading Test with Over One Hundred Thousand Children. *Jr. Ed. Ps.* 9: December, 1917, pp. 600-608.

*THE STARCH SILENT READING TESTS, 1914

(University of Wisconsin)

This test consists of eight selections printed on separate sheets. Each test is intended for a different grade, No. 1 for the first grade, No. 2 for the second grade, etc. Directions for giving and scoring the tests are included. The pupils are told to read silently, as rapidly as they can, and at the same time to grasp as much as they can. They are also told that they will be asked to write down in their own words as much as they remember of what they read. They are told, furthermore, not to read anything over again but to read on continuously. The blanks are distributed face down, and at the signal "turn and start" they begin reading. The time limit for the test is exactly 30 seconds. The pupils draw a line around the last word read to indicate how far they have read. The pupils are then told to turn the blanks over and write down on the back all they remember of what was read. There is no time limit for reproducing the story, but the teacher should be careful

that they do not copy from each other. The speed of reading is determined by ascertaining the number of words read per second. To find the score for comprehension the written account is carefully read and all incorrect or irrelevant words, as well as repetitions, are crossed out. The remaining words are counted and the number is taken as the score for comprehension. The average number of words which have to be discarded by this method is about 7%, so that if only the class averages are desired the words in the reproduction can simply be counted and 7% deducted. Of course this method cannot be applied if the results of the individual pupils are desired. The author recommends that a second test be given, using the test sheet for the grade below, and that the average of the two results be taken as the final scores for both speed and comprehension. Standards are given on the direction sheet.

This test seems to be a good one, though the plan of counting the number of words in the reproduction of the pupil in order to obtain a score for comprehension has been questioned. Another possible disadvantage is that some of the selections used are familiar to some classes. The advantage is that there is a different selection to be read by each grade. The print of the selections for lower grades is larger than that for the upper grades, and the selections are easier.

Address: Dr. Daniel Starch, University of Wisconsin, Madison, Wisconsin.

References: Starch Educational Measurements, Macmillan and Company, New York City.

Starch, D.—The Measurement of Efficiency in Reading. Jr. Ed. Ps. 6: 1-24, January, 1915.

***THORNDIKE READING SCALE ALPHA (1914) AND ALPHA 2 (1915) FOR MEASURING THE UNDERSTANDING OF SENTENCES**
(Teachers' College, Columbia University)

Since scale Alpha 2 is only an improvement of the earlier scale Alpha, only scale Alpha 2 will be described.

This scale is published in two parts, part one for grades 3 to 5, part two for grades 6 to 12. The two scales overlap to some extent. Part one consists of 4 sets of reading selections with questions added. Each set has a score value attached and the set is marked "passed" if the student succeeds in answering 80% of the questions correctly. Part two consists of 4 sets of paragraphs, beginning with set IV and extending to set VIII. These are more difficult than those for part one, and consequently the score values are higher. A score key containing the correct answers to the questions and a correction table for interpolating the scores and class record sheets containing space for the tabulation of the answer to each question of each paragraph are included. Directions for giving and scoring the tests and standard medians are given in a pamphlet, "Directions for Thorndike Reading Scale Alpha 2," published by Teachers' College, Columbia University, New York City.

Twenty-five minutes are allowed for the test. The test is an excellent one, though it is not, strictly speaking, a timed test, since 25 minutes is usually long enough to complete the work.

Address: Bureau of Publications, Teachers' College, Columbia University, New York City.

References: Thorndike—An Improved Scale for Measuring Ability in Reading, Teachers' College Record, 16: November, 1915, pp. 31-53; and 17: January, 1916, pp. 40-67.

Thorndike—Reading as Reasoning, A Study of Mistakes in Paragraph Reading, Journal of Educational Psychology, 8: June, 1917, pp. 323-332.

Kelley, T. L.—Thorndike's Reading Scale. Alpha 2 Adapted to Individual Tests. Teachers' College Record 18: May, 1917, 253-260.

THE MINNESOTA READING TEST: UNDERSTANDING OF SENTENCES, SCALE BETA I

(Date not obtainable)

There are two scales in this series, the first for grades 3 to 5, the second for grades 6 to 9. The scale is a modification and extension of Thorndike's

Scale, Alpha 2. It is put up in a little more convenient form than the Thorndike Scale.

The test for grades 3 to 5 consists of 5 paragraphs to be read and questions to be answered on each paragraph. The respective scores of 10, 30, 40, 50 and 60 are assigned to the paragraphs. On the back of the folder is a preliminary test which contains three sentences and three questions under each sentence. The child is to read each sentence and then answer the three questions. After the preliminary test is given the directions for the test proper are read, and the children read the paragraphs and answer the questions in writing.

The score for the individual pupil is the "highest numbered line which the child does with one or no omission or error." The class score is the number of the line in which the percentage of error is 20. A table is included to aid in interpolating the scores. There is also a key or answer sheet for determining the answers. A folder of complete directions for giving and scoring the test is included with the test material.

The test for grades 6 to 9 is similar to the one just described. This test begins with value 40 and extends to value 70. Qualities 40, 50 and 60, which are given for the former test, are repeated here. On the first page of both of these tests is an individual score sheet. The scale is issued in two forms which are approximately equal in value, so that the test may be repeated by the same students. Record sheet number 2 should be ordered with the test for recording the results. This test has a high reliability co-efficient. It is not completely standardized, however. Tentative standards can be procured from the following address.

Address: University of Minnesota, Minneapolis, Minnesota.

Reference: Haggerty, M. E.—The Ability to Read: Its Measurement and Some Factors Conditioning It. Indiana University Studies 4: January, 1917, Bloomington.

***THE FORDYCE SCALE FOR MEASURING THE ACHIEVEMENTS IN READING, 1916**

(University of Nebraska)

This test consists of 2 selections, one for grades 3 to 5, and the other for grades 6 to 9. The test for the lower grades is called "Narcissus" and is taken from the Greek legend. The test for the upper grades is the "Spirit of Spring." Test number one, the story of Narcissus, contains 300 words, and test number two contains 512 words. One minute is allowed for giving the test, and the pupil is asked to make a mark around the word he is reading when time is called. For test number one the percentile grade for rate is found by dividing the number of words read in the given time by three, since there are 300 words in the selections. For test number two the percentile grade for rate is found by dividing the number of words written by five. The number of words can be counted easily by referring to the key given in the leaflet of instructions.

After the speed test is completed the children are to read the remainder of the selection so that they can answer the questions on it. The blanks containing the questions on the passages are then passed around for the purpose of testing the quality of reading. Ten minutes are allowed for answering the questions on test number one, and 15 minutes for test number two. Each one of these questions is evaluated, and the value of each question is given in the booklet of instructions. The total values equal 100. A score card or key is also provided in the booklet of instructions, which gives the correct answer for each question. The values for test number two, "The Spirit of Spring," have been revised recently. Sheets containing the values for this test, and also new standards, can be secured from the address below.

The test includes some practice exercises for the benefit of the students who are behind standard, especially in rate. These exercises consist of groups of words arranged in 4 columns. The first column contains 24 combinations of 2 words each, exercise II, 24 groups of 3 words each, exercise III, 24 groups of 4 words each, and exercise IV, a selection, "The House That Jack Built," printed in groups of from three to five words in a group. The practice is con-

ducted by having the pupil read and grasp groups of words as rapidly as possible. He begins with the simpler exercises and proceeds to the more difficult ones. The author finds that the slow readers are usually those who have to read in small units, and that the perceptual span increases with the ocular span, for "the mind passes more rapidly to meaning from phrases or sentences than from single words." The author asserts that through ten minutes' practice daily of similar exercises he was able to double his rate of reading within one year's time.

Address: Dean Charles Fordyce, University of Nebraska, Lincoln, Nebraska.

Reference: Fordyce—A Scale for Measuring the Achievements in Reading, same address.

***MONROE'S STANDARDIZED SILENT READING TEST, 1917**
(University of Illinois)

This test is devised on the same principle as the Kansas Silent Reading Test, but is an improvement over this earlier test. The test is put up in exactly the same form. The directions are included on the first page of the folder and instructions are given which are to be read aloud by the children and teacher in concert. An illustrative example is also given on the first page and the children respond to this test for preliminary practice. There are three tests in the series, test 1 for grades 3, 4 and 5; test 2 for grades 6, 7 and 8; and test 3 for the high school.

Test 1 consists of 16 paragraphs which are to be read silently. One question is asked on each paragraph to which the child must respond, usually by adding a word in the blank space or by underlining some word given in a series. The response is always very short and definite. Each paragraph has in the left margin of the page a rate value attached and in the right margin of the page a comprehension value. Thus the pupils are scored for rate and comprehension separately. The score for rate is the sum of the rate values of the questions answered and the score for comprehension is the sum of the comprehension values of those paragraphs responded to correctly. An answer sheet is included on the back page of the record blank to aid in scoring the results. Standards are also given on the record blank, as well as instructions for finding the median and tabulating the results. The standards are revised from time to time so that the latest standards should be requested when ordering the tests. Five minutes are allowed for the tests. Directions for scoring the tests and finding the median and recording the results are given on the record sheet. Test 2 is similar except the questions are a little more difficult and there are only 14 paragraphs in this test. Twelve paragraphs are included in test 3.

This test is an improvement over the Kansas Silent Reading Test in two respects. In the first place the paragraphs are selected from readers and are very much less puzzling in their nature than those of the Kansas Test. They are more typical of the paragraphs which the pupil finds in his general reading. The responses too are simpler and more natural than those in the Kansas Test, so that this test seems to be less a test of general intelligence and more nearly a fair test of reading ability than the Kansas Test. In the second place there is a separate score for rate and for comprehension, while both of these factors are included in the score for comprehension in the Kansas Test. A disadvantage is that this test has not been so widely standardized as the Kansas Test and has not been so often used in the surveys, so that the means of comparison with achievements of other cities is somewhat more restricted. This test is employed by the State Department of Public Instruction of Colorado.

Address: Kansas State Normal School, Emporia, Kansas.

References: Monroe, W. S.—Monroe's Standardized Silent Reading Test. Jr. Ed. Ps. 9: 303-312, June, 1918.

Witham, E. C.—Scoring the Monroe Silent Reading Test. Jr. Ed. Ps. 1918: 9, 516-518.

***THE COURTIS STANDARD RESEARCH TESTS, SERIES R, SILENT
READING TEST NO. 2, 1917**
(Bureau of Research, Detroit)

This test is intended for grades 2-5 inclusive. The results of these tests when given to the upper grades have proved to be unreliable and for that reason Curtis suggests that in grades 6-8 the Monroe Silent Reading Test be substituted for this one. This test consists of a story of two pages in length containing 567 words. This part of the test is used as a speed test and the pupils are instructed to make a mark around a word which they are reading at the end of 30-second intervals which are called out by the teacher during the test. The speed test lasts exactly three minutes. This constitutes part 1 of the test.

Part 2 of the test consists of the same selection divided into 14 paragraphs. There are 5 questions asked under each paragraph, all of which can be answered by "yes" or "no." The pupils are asked to read these selections one at a time and then to answer the 5 questions by the words "yes" or "no." The pupils are allowed to reread these paragraphs in order to get the correct answers. They are also warned not to guess at the answers. The pupils are asked to draw a circle around the number of the question they are answering at the end of each minute when the signal is given by the teacher. The rate score is the number of words read per minute in the speed test. The score for the questions is the number of questions answered in 5 minutes. In case a pupil finishes the test before five minutes the number of questions answered in 5 minutes is computed for the average number answered per minute.

In addition to this another score called the index of comprehension must be found. This is done by subtracting the wrong answers from the right answers algebraically and then dividing the difference by the number of right answers. A table is also given for finding the index of comprehension. Standards are given on the record sheet. A graph sheet is also included for drawing a graph of the class record. Also an individual record card is provided for each pupil with answers printed on the back. This card provides separate space for recording all of the results in detail. Also a very convenient class record sheet is provided as an aid to tabulating the results. This sheet also contains standards for the test. Three folders of instructions are also included with the test. Folder A contains general instructions to examiners in all subjects; folder B-R contains detailed instructions for giving and scoring the tests; folder D-R contains instructions for completing the scoring, for making records, and interpreting the results. The instructions are very complete. The test itself is put up in a very convenient folder.

Some of the possible disadvantages of the test are that the pupils merely reply by the words "yes" or "no" and consequently there is a possibility of guessing. The index of comprehension, however, offers opportunity to check up on this. Also only questions are asked on the contents of the page and no opportunity is given for reproducing the selection.

Address: S. A. Curtis, 82 Eliot Street, Detroit, Michigan.

Reference: Curtis—The Problems of Measuring Ability in Silent Reading. American School Board Journal 54: May, 1917, pp. 17-18, and p. 81.

***THE HOLMES READING TEST (Date Uncertain)**
(Harvard University)

There are two forms of this test, each consisting of a reading selection, the first selection, "The Rich Man," containing 753 words, the second selection, "The Great King," containing 786 words. Twenty seconds are allowed for reading as much of the material as possible. When the time is called to stop, the pupils are to underline the last word read. The pupils are told to finish reading the story, however, so that they may know the contents of it. The children then reproduce the story as accurately as possible. Then a list of standardized questions is copied on the blackboard for the children to answer.

The scoring for rate is done by computing the number of words read per minute. A score sheet is provided with the principal ideas of the test underlined. The score for comprehension is found by counting the number of ideas indicated by the key which are contained in the student's reproduction of the story. The number of ideas correctly reproduced is multiplied by $2\frac{6}{7}$ in order to determine the final score for the reproductions. The answers to the questions are weighted and the score values indicated opposite the questions. There is also an answer key for these questions which indicates percentages to be given for each part of the answer. Another score sheet is provided which has the words in the selections numbered so as to obviate the necessity of counting the words to determine the score for rate. Complete directions for giving and scoring the test are added.

The method of scoring this test for comprehension is extremely good. The test was used in the survey of the schools of Brookline, Massachusetts, and these results may be used for comparison.

Address: Harvard University, Cambridge, Massachusetts.

Reference: Survey of the Brookline Schools, Brookline, Massachusetts.

THE SACKETT READING TEST, 1919
(University of Texas)

This test consists of a story of 110 words. The directions instruct the children to read the story through till they are sure they know it well. When they have finished reading it they write the story in their own words in the space provided below. On the back of the sheet are 15 questions to be answered on the context. Some of the questions are of the nature of performance tests. Standards are available from the third grade to the university Sophomore Class.

Address: L. W. Sackett, University of Texas, Austin.

***THE HAGGERTY ACHIEVEMENT EXAMINATION IN READING,**
SIGMA 1, FOR GRADES 1-3, 1920
(University of Minnesota)

Test 2 of this series, which is given first, consists of 20 questions which are followed by the words "no" and "yes." The response is made by underlining the correct answer. The questions are increasingly long and difficult. The first, for example, is "Can you eat?" The last is "Do convicts sometimes escape from prison?"

Test 1 consists of reading sentences or paragraphs, and responses to be made or questions to be answered. The simplest ones are printed in very large type and are of the nature of performance tests, for example: "Put a tail on this pig." A sketch of the pig without a tail appears on the other side of the page, and the child is to draw the tail. Test 1 contains 25 questions or responses. Questions 1-8 are in 16 point type, 9-21 in 14 point type, and 22-25 in 12 point type. The test appears in booklet form which is well illustrated by pictures. These serve both as incitements to interest and as means of making responses. The reading paragraphs become increasingly long and difficult. The questions are usually very simple, so as to make the test really a test of reading rather than a test of intelligence. This test was used in the recent Virginia School Survey. Directions for giving and scoring the test are included in a separate manual. A score card, record blanks, and standards accompany the tests.

Address: M. E. Haggerty, University of Minnesota, Minneapolis, Minnesota.

References: Haggerty, M. E.—Manual of Directions, World Book Co., Yonkers, New York, 1920.

Virginia School Survey, Richmond, Virginia.

THE ADAMS SILENT READING TESTS, 1920
(State Normal School, Plymouth, New Hampshire)

This test consists of 8 selections, one for each grade. The first is printed in large, bold type, the second, third and fourth in 12 point type, and the others in smaller type. The reading material is chosen with a view to the

pupil's pedagogical needs, the developmental stage of the child, the child's interests and correlation with school activities.

The speed test is given by allowing the children to read for 30 seconds and then having them draw a line around the last word read, and is expressed as the number of words read per minute. The comprehension score is determined by the answers to 10 questions on the back of the test sheet. To secure the comprehension score each correct answer is marked "1." and the results are multiplied by the grade of the test so that the highest possible score for comprehension in the first grade is 10 and for the eighth grade 80. The pupils are also allowed to report the story and the teacher may check up by the questions. The questions may be answered orally in the first three grades.

Score cards and a direction sheet for giving and scoring the tests are included. The chief advantage of the test is its simplicity and the adaptation of the reading material to the interests of the child.

Address: Edward E. Babb and Company, 93 Federal Street, Boston, Mass.

2. Vocabulary Tests

*THE THORNDIKE READING SCALES A, A-2 AND B: WORD KNOWLEDGE OR VISUAL VOCABULARY, 1914-1915

(Teachers' College, Columbia University)

The latter two are modifications of Scale A, which was published in 1914. Scale A contains fewer words than the later editions. Scale A-2, x series, will serve as an example. Other series are published as equivalent tests. Scale A-2 consists of 13 lists of words arranged in so many lines and in increasing order of difficulty, each of which has the value indicated in the margin. Above these words are eight-sets of directions. The first set of directions is to write the letter "F" over every word on the page that means a flower. The second is to write the letter "A" over every word on the page that means an animal. There are six other sets of directions of a similar nature. The pupil's score is the score value of the most difficult line of which the pupil answers 8 of the 10 words in the line correctly.

Address: Bureau of Publications, Teachers' College, Columbia University, New York City.

References: Thorndike—Measurement of Achievement in Reading: Word Knowledge; Teachers' College Record 17: November, 1916, pp. 430-454.

Thorndike—Measurement of Ability in Reading Preliminary Scales and Tests. Teachers' College Record 15: September, 1914.

Childs, H. G.—A Half-Year's Progress in the Achievement of One School System as Measured by the Thorndike Visual Vocabulary Test. National Society for the Study of Ed. 15th Yearbook, 1916, Part I, 79-83.

THE MINNESOTA READING TEST: VISUAL VOCABULARY SCALE R-2

Form 4 for Grades 3 and 4

Form 5 for Grades 5, 6, 7 and 8

(Date Unknown)

(N. E. Haggerty, University of Minnesota)

This scale is a modification of the Thorndike Scale just described. For Form 4 there are 4 lines, each containing 5 words. There are 12 sets of directions at the top. An illustrative preliminary test is given on the back of the sheet. Form 5 is arranged in the same manner but contains 6 lines of 5 words each and 15 sets of directions. These tests differ from the Thorndike Test chiefly in the use of fewer words, a different selection of words, and a greater number of sets of directions, requiring a greater number of responses.

Address: Bureau of Co-operative Research, University of Minnesota, Minneapolis, Minnesota.

Reference: Haggerty, M. E.—Scales for Reading Vocabulary of Primary Children. *El. Sch. Jr.* 17: 106-115, October, 1916.

***THE STARCH ENGLISH VOCABULARY TEST, 1916**
(University of Wisconsin)

This test consists of two lists of 100 words each. The words are chosen from the dictionary by the method of random sampling, after the technical words are eliminated. The child is to check off the words whose meaning he is sure of and can use correctly, and write the meaning after the words with which he is familiar but of whose meaning he is not quite sure. The children are also told that they will be asked by the examiner to write the meaning after any of the difficult words that they may have checked. This is to insure the examiner that the child knows the meaning of the words checked off. Directions and standards accompany the test. The words of the test are not those comprising the vocabulary of children. The test is not a measurement of words needed by the child, but it is of value since the score indicates the percentage of non-technical words of the English language that the child knows.

Address: Dr. Daniel Starch, University of Wisconsin, Madison, Wisconsin.

References: Starch—The Measurement of Efficiency in Reading, *Journal of Educational Psychology* 6: January, 1915, pp. 1-24.

Starch—Educational Measurements, Second Edition, Macmillan and Company, 1920, New York City.

THE SOUTHINGTON-PLYMOUTH ENGLISH VOCABULARY SCALE, 1919

This test consists of 50 rather common words arranged in the order of increasing difficulty, and numbered. On the other side of the page are definitions for these words. The pupil must place before the definition the number of the word which is defined.

Address: Supt. Witham, Southington, Conn.

3. Oral Reading Tests

***THE GRAY STANDARDIZED READING PARAGRAPHS, 1914**
(William S. Gray, Chicago University)

This test consists of 12 reading paragraphs of about the same length. The first ones are very simple and the others are increasingly difficult. The first three selections are printed in large type suited to beginning pupils, and the remainder of the paragraphs are in smaller but sufficiently large and clear type. Each pupil is tested separately.

The time is taken when the pupil begins reading each paragraph and again when the pupil finishes reading the paragraph. The errors in pronunciation and of omissions, substitutions, and repetitions are marked in the test, and the time required for reading the paragraph is recorded on the margin of the test sheet. Each pupil is allowed to continue reading until he makes 7 errors in each of 2 paragraphs. The pupil's score varies inversely with the time required for reading the paragraphs and the number of errors made.

Complete directions for giving the test and scoring and tabulating the results are given on the back of the score sheet for reading.

This test is probably the best test of oral reading for all the grades. It was used in the survey of the Cleveland schools and in several other surveys, including the survey of the schools of Sterling, Colorado, so that standards and means of comparison with other school systems are available.

Address: William S. Gray, School of Education, University of Chicago, Chicago, Illinois.

Reference: Gray—Methods of Testing Reading, *Elementary School Journal* 16: January, 1916, pp. 231-246, and February, 1916, pp. 281-298.

***THE JONES SCALE FOR TEACHING AND TESTING ELEMENTARY READING, 1915**

This is a test of oral reading for the first three grades. It is made up of all sight words and of phonograms recurring ten times or more in ten

widely used primers. There are really two tests, one for sight words of which there are 192, and one for the phonetic test containing 118 words. The words are weighted according to the frequency of their occurrence in the primers. The score for a pupil is the percentage which the sum of the values attached to the words correctly pronounced is of the sum of the values of all the words. The total value of all the phonograms is 15,657; that of the sight test is 17,565. To find the score for a class of ten pupils selected at random from a larger class, the sum of the total values of the words missed by the entire ten pupils is subtracted from the above number multiplied by ten. The difference is divided by 156,570 for the phonograms, or 175,650 for the sight words, so that the resulting score is expressed in percentage. Standards are given with the test materials. Complete directions for giving and scoring the test and teachers' check cards with the weighted values of the words and space for the individual records of the pupils, also pupils cards for reading the words accompany the tests. There is a separate sheet containing standards from 20 schools. A monograph by the author entitled, "Standard in Mechanics by Elementary Reading," goes with the test.

This test is a very good one for vocabularies of young children, although it takes no consideration of speed reading and errors of insertion, omission, repetition and such mistakes. It is intended only as a test of recognition and vocabulary and pronunciation of unrelated words. Another objection which has been advanced with regard to these tests concerns the method by which the words were evaluated. The fact that "the" occurs a great many more times in the primers than the word "pieces" may not necessarily entitle it to a proportionately high value. The test is, however, a great aid to the teacher.

Address: R. G. Jones, 1453 Marlowe Avenue, Lakewood, Ohio.

References: Jones—Scale for Teaching and Testing Elementary Reading, Rockford, Illinois.

Jones—Standard in Mechanics of Elementary Reading, Lakewood, Ohio.

THE MINNESOTA VISUAL VOCABULARY TEST FOR GRADES 1 AND 2 (Date unknown. Haggerty, University of Minneapolis)

This test is a simplification of the above test by Jones. The test consists of two pupil's cards, form 1 containing 30 sight words and form 2 containing 25 phonetic words. The words are arranged in order of increasing difficulty, 5 words in a group or line.

"The child's score is the highest line in which he makes not to exceed 1 error or omission. The class score is the number of the line in which the per cent of error is nearest 20."

Two class record sheets with space for recording the results of 25 pupils are supplied, one for the sight test and one for the phonetic test. A sheet of directions for giving and scoring the test is also provided.

Address: Bureau of Co-operative Research, Minneapolis, Minnesota.

Reference: Haggerty—Scales for Reading Vocabulary for Primary Children, *Elementary School Journal* 17: October, 1916, pp. 106-115.

THE PRICE PRACTICAL ORAL READING TEST FOR GRADES 2 TO 8 **OCTOBER, 1916** (Superintendent, Enid, Oklahoma)

The same general plan is followed in the test for each grade. The test for each grade consists of two selections to be read, test number 1 to be given early in the school year and test number 2 near the close of the school year. Each test consists of an interesting story or selection. The number of words in each line is indicated in the right margin. The children are tested individually for one minute each and the mistakes are recorded for: 1, words miscalled; 2, words put in; 3, words left out; 4, transposed words. Two persons are required to give the tests, one to give the directions and to keep time, and the other to record the number and nature of the mistakes and count the number of words read. The score for each pupil is found "by multi-

plying the per cent of words read correctly in the total number of words read, by the number of words read correctly. The class score is the approximate median of the individual scores.

Directions for giving the test are printed on the first page of the folder, and directions for scoring the test, as well as standards based on the records of about 200 children of each grade, are given on the back of the score sheet. The score sheet contains columns for age, pupil's name, words miscalled, words put in, words left out, transposed words, number of words read, number of mistakes, and score. The standards are tentative, but are to be revised from time to time. The tests have the advantage of providing different selections for every grade so that the reading material is suitable for the grade intended. The principle of the test seems to be a good one.

Address: E. D. Price, Enid, Oklahoma.

GENERAL BIBLIOGRAPHY ON READING TESTS

- Anderson, H. W.—Measuring Primary Reading in the Dubuque Schools. The Harris-Anderson Test. Dubuque, Iowa, 1916, 23 pp.
- Bobbitt, F.—Reading in the Elementary Schools of Indianapolis, V and VI: the Reading Materials. *El. Sch. Jr.* 19: 665-688, 741-761.
- Boston. Dept. of Ed. Investigation and Mes.—Standards in Silent Reading, with Suggestions on How Teachers May Test Their Pupils in Silent Reading. Boston, 1916, School Document No. 18, Bul. No. 12, 24 pp.
- Brown, H. A.—The Significance of the Measurement of Ability to Read. *Education* 36: May, 1916, 589-610.
- Gray, W. S.—Methods of Testing Reading. *El. Sch. Jr.* 16: January-February, 1916, 231-246, 281-298.
- Gray, W. S.—The Use of Tests in Improving Instruction. *El. Sch. Jr.* 19: 1918, 121-42.
- Gray, C. T.—Types of Reading Ability as Exhibited through Tests and Laboratory Experiments. Chicago University Press, 1917, 196 pp.
- Gray, W. S.—Selected Bibliography upon Practical Tests of Reading Ability. Fourteenth Yearbook, National Society, Pt. I, 1915, 59-60.
- Gray, W. S.—A Study of the Emphasis on Various Phases of Reading Instruction in Two Cities. *El. Sch. Jr.* 17: November, 1916, 178-186.
- Gray, W. S.—The Relation of Silent Reading to Economy in Education. Sixteenth Yearbook, National Society, Pt. I, 1917, 17-32.
- Gray, W. S.—Value of Informal Tests of Reading Accomplishment. *Jr. Ed. Research* 1: February, 1920, 103-111.
- Haggerty, M. E.—The Ability to Read: Its Measurement and Some Factors Conditioning It. *Indiana University Studies* Vol. 4, No. 34, January, 1917. Bloomington, Indiana.
- Hayes, S. P.—Report of Preliminary Tests in Reading. *Publ. Penn. Institution for Instruction of the Blind*, 1: 1918, pp. 20.
- Heilman, J. D.—A Study of the Mechanics of Reading. *Colorado State Teachers Bul.* 1918, Greeley.
- Hill, D. S.—Experimental Studies of Achievement in Reading. New Orleans, La., Public Schools, Annual Report 1915-16, Pt. 3, 61-72.
- Judd, C. H.—Demonstration of the Gray Reading Test. Also Reading Tests. Fourth Conf. on Educational Mes. 1917, Indiana University, Bloomington.
- Judd, C. H.—Report of Committee on Tests and Standards of Efficiency in Schools and School Systems: Reading Tests. *Proc. N. E. A.* 1915: 561-565.
- Judd, C. H.—Reading. Fifteenth Yearbook, National Soc. Pt. I, 1916, 111-119.
- King, I.—A Comparison of Slow and Rapid Readers. *Sch. & Soc.* 4: November 25, 1916, 830-834.
- King, I.—Comparison of the Efficiency of Slow and Rapid Readers. *Sch. & Soc.* 6: August 18, 1917, 203-204.
- Mead, C. D.—Silent Versus Oral Reading with One Hundred Sixth-Grade Children. *Jr. Ed. Ps.* 6: June, 1915, 345-348.

- Merrill, Maud—A Scale for the Individual Measurement of Reading Ability. Jr. Ed. Ps. 10: October, 1919, 389-400.
- Otis, A. S.—Considerations Concerning the Making of a Scale for the Measurement of Reading Ability. Ped. Sem. 23: December, 1916, 528-549.
- Peters, C. C.—The Influence of Speed Drills upon the Rate and Effectiveness of Silent Reading. Jr. Ed. Ps. 8: June, 1917, 350-356.
- Pintner, R., and Gilliland, A. R.—Oral and Silent Reading. Jr. Ed. Ps. 7: April, 1916, 201-212.
- Richards, O. M., and Davidson, T. E.—Correlations of Single Measurements of Some Representative Reading Tests. Sch. & Soc. 4: September 2, 1916, 375-377.
- Schmidt, W. A.—An Experimental Study in the Psychology of Reading. Chicago University Press, 1917, pp. 126.
- Starch, D.—The Reliability of Reading Tests. Sch. & Soc. 8: 1918, 86-90.
- Thorndike, E. L.—Tests of Oral and Silent Reading. Indiana University Bul. 12, 1914, 37-51.
- Uhl, W. L.—The Use of the Results of Reading Tests as a Basis for Planning Remedial Work. El. Sch. Jr. 17: December, 1916, 266-275.
- Wassen, A. W.—A Report of an Experiment in the Use of the Kansas Silent Reading Test with Korean Students. Ed. Adm. & Sup. 3: February, 1917, 98-101.
- Whipple, G. M., and Curtis, J. N.—Preliminary Investigation of Skimming in Reading. Jr. Ed. Ps. 8: June, 1917, 333-349.
- Ziedler, R.—Tests in Silent Reading in the Rural Schools of Santa Clara County, California. El. Sch. Jr. September, 1916, 65-62.

VIII—GEOGRAPHY

GEOGRAPHY TESTS, 1915

The difficulties of testing geography are due chiefly to unformulated aims, purposes, and value of geography instruction, and to lack of analysis of the subject. Possibly most teachers teach the mere facts of geography with the stress on definitions and formal knowledge. Others emphasize the location of important centers, still others, products from different regions, and again others, commercial geography, and still others, physical geography. Other teachers emphasize national geography; some, the influence of climate and geographical relations upon peoples; some, political geography, or various other elements.

It is difficult to test geographical appreciation or geographical thinking, so that most of the standard tests now available must confine themselves largely to an exploration of the facts in the common knowledge of the pupils.

THOMPSON'S MINIMUM ESSENTIALS IN GEOGRAPHY, 1908

These consist of two sets of tests, one an oral drill on answering questions of geography and the other the written test. These tests were intended originally chiefly as aids in teaching rather than as standardized tests, but have been found useful also for the latter purpose.

Address: Ginn and Company, New York.

THE BOSTON GEOGRAPHY TESTS, JANUARY, 1915

These tests consist of two sheets, one for the geography of the United States and one for the geography of Europe.

The map of the United States occupies the upper part of the first sheet. Below it are blanks for the name, age, sex, school and grade of the pupil. Then follow 10 questions on the geography of the United States. Cities mentioned are to be located on the map. Products for which the cities are noted are to be given. Localities which produce certain staple products and reasons for growth of certain large cities are to be given. Several questions on climatology, and one question on the commerce of San Francisco and New York compared are given.

The test on the geography of Europe consists of a map of Europe and seven questions on European geography. These questions are:

1. Locate on the map 2 seaports of European Russia.
2. Why are the seaports of Russia not so important as the seaports of England?
3. Of what value to the countries of Europe are their colonies in other parts of the world?
4. Why does England import large quantities of wheat?
5. Write on the map the names of the leading manufacturing countries of Europe.
6. Why has Germany become very important as a manufacturing country?
7. Why is the climate of Italy different from that of Germany?

These questions teach many of the important aspects of European geography. The disadvantage of these tests lies in part in their brevity. They were not printed for distribution. They were used, however, in the survey of the Boston schools, and the scores from this survey are available for comparison.

Address: Department of Educational Investigation and Measurement, Boston, Massachusetts.

Reference: Geography: Bulletin No. 5, School Document No. 14, 1915, Boston, Massachusetts, Dept. of Ed. Investigation and Mes.

THE BUCKINGHAM GEOGRAPHY TESTS, JANUARY, 1916 (University of Illinois)

This test was used in the survey of the Gary and Prevocational Schools of New York City. There are 20 questions in the test and these have been evaluated experimentally, and standardized. The tests are not printed for distribution but are given in full in the following reference.

Address: H. R. Buckingham, University of Illinois, Urbana, Illinois.

Reference: Buckingham—A Survey of the Gary and Prevocational Schools of New York City.

THE STARCH GEOGRAPHY TEST, SERIES A, 1917 (University of Wisconsin)

This is a test containing 78 questions and statements arranged in the form of mutilated sentences. The blanks are to be filled in with the correct responses. There are several blanks on each question, and the score consists of the number of blanks correctly filled in. Directions for giving and scoring the test are included in the folder. Standards are given for grades 5 to 8, and were obtained from approximately 1,300 pupils. A score key should be procured to aid in scoring the test.

This test is very complete, but it has the fault of emphasizing disconnected facts and formal knowledge, much of which is less important than geographical thinking, reasoning and judgment. The test is not a timed test.

Address: Dr. Daniel Starch, University of Wisconsin.

***THE HAHN-LACKEY GEOGRAPHY SCALE, 1917, SECOND EDITION, 1919** (State Normal School, Wayne, Nebraska)

This geography scale is arranged on the same principle as the Ayres Spelling Scale. The scale is based upon 283,100 answers by 1,696 pupils in 12 schools. The scale consists of 216 questions arranged in 25 columns lettered from A to Y. The questions in each column are of an approximately equal degree of difficulty. In no case does the absolute value of an exercise differ from the approximate value by more than 00.4%, and this only in a very few cases. The scale is so arranged that from 15 to 20 tests can be given to each grade without the repetition of an exercise.

The percentage of correct answers to be expected from each school grade for the questions in each column is indicated at the top of the scale, so that these percentages serve the purpose of standards, as is the case in the Ayres Spelling Scale. To give the test, several of the questions, usually ten,

are written on the board as in an ordinary examination, and the pupils are told to answer these in writing. The scale is intended for grades 4 to 8 inclusive.

This is one of the best geography scales. A disadvantage, however, is that no map work is included, and many of the questions are somewhat formal in character.

Address: H. H. Hahn and E. E. Lackey, Wayne State Normal, Wayne, Nebraska.

References: Hahn and Lackey—Monograph Describing the Geography Scale, Wayne State Normal, Wayne, Nebraska.

Lackey, E. E.—Measuring the Ability of Children in Geography, *Journal of Geography* 16: 184-188, January, 1918.

Lackey, E. E.—A Scale for Measuring the Ability of Children in Geography, *Journal of Educational Psychology* 9: 443-451, October, 1918.

Matthewson, C. A.—Some Results with the Hahn-Lackey Scale in Geography, *Journal of Educational Psychology* 9: 581-587, December, 1918.

THE COURTIS SUPERVISORY TEST IN GEOGRAPHY, TEST A FORM A, 1918

(Bureau of Research, Detroit)

This is intended as a test of the knowledge in only one small field of geographical study, the location of states and of prominent cities. This test consists of a map of the United States, on which each of the 48 states is numbered. Below the map is a list of the states and after the name of each state is to be placed its number on the map. The second part of the test, the location of cities, consists of locating by number states in which 30 prominent cities are located. The time allowed for the first test is four minutes, that for the location of cities, two minutes. The scores are the number of states and the number of cities correctly located. Answer cards, individual record cards, class record cards, and a large class record sheet arranged for filing are included. Standards may be obtained from the following address.

Address: S. A. Courtis, 82 Eliot Street, Detroit, Michigan.

Reference: Courtis, S. A.—Measuring the Effects of Supervision in Geography, *School and Soc.* 10: 61-70, July 19, 1919.

*WITHAM'S STANDARD GEOGRAPHY TESTS, 1918

(Superintendent, Southington, Connecticut)

There are several of these tests, one on the United States (1918), one on South America (1919), and one on the world. The test on the world is to be given to fifth grade pupils. The test on the United States is intended for sixth grade pupils, and that of South America for the more advanced pupils. The test on the United States will be described as a typical example. Following are the questions:

- I. Draw on the space below an outline of the United States.
- II. On the map just drawn, write the names of what bounds the United States on all sides.
- III. Draw an outline on the above map of the state in which you live. Locate the capital of your state by means of a small circle. Write its name.
(On the second page of the leaflet is a printed outline map of the United States.)
- IV. On the printed outline map on the opposite page, neatly letter the names of the following features. You may use abbreviations. (Then follow the names of ten important rivers.)
- V. Locate the following mountains. (The names of five mountain ranges follow.)
- VI. Locate the following five lakes and gulfs.
- VII. Locate on the outline map the list of fifteen cities.
- VIII. Locate on the map a list of five states. On the last page of the folder are ten maps showing the industrial regions of the United States. The names of these industrial regions are to be found at the top of the page.

IX. Identify on the map each of the ten industrial regions.

A folder of directions for giving and scoring the tests and recording the results, and also directions for drawing a graph of the results are included. On the other side is a class record sheet, and a blank graph for the class scores on all of the questions. The tests on the world and on South America follow the same plan. Standards are available for all of these tests.

This test has been criticised because of the predominance of questions on locations of places, and the emphasis on facts. Most of these elements should, no doubt, be known to one familiar with geography, since they are all rather important questions. The last question in the test described is important because of the knowledge of commercial and industrial geography involved. Many other phases of geography, however, are not included, and so the test is not entirely complete.

Address: Ernest C. Witham, Southington, Connecticut.

References: Witham, E. C.—A Minimum Standard for Measuring Geography. *Amer. Sch. Bd. Jr.* 50: 13-14, January, 1915.

Witham—Standard Geography Test—the Word. For Fifth Grades. *Jr. Ed. Ps.* 9: 432-442, October, 1918.

THE BUCKINGHAM GEOGRAPHY TEST FOR GRADES 7 AND 8, 1920
(University of Illinois)

This test is in preparation at the present time. It may be secured from the Bureau of Educational Research, University of Illinois, Urbana, Illinois.

BIBLIOGRAPHY ON TESTS IN GEOGRAPHY

Bagley, W. C.—The Determination of Minimum Essentials in Elementary Geography and History. *Fourteenth Yearbook National Society*, 1915, Pt. 1, 131-146.

Branom, M. E., and Reavis, W. C.—The Determination and Measurement of the Minimal Essentials of Elementary School Geography. *Seventeenth Yearbook National Society*, Pt. 1, 27-39.

Taylor, E. G. R.—Children's Mistakes in Geography. *Jr. Ed.* 50: 1918, 322-323.

Whitbeck, R. H.—A Test in General Geography. *Jr. Geogr.* 16: December, 1917, 149-152.

IX—HISTORY

HISTORY TESTS, 1916

The difficulties in devising history tests are similar to those encountered in the geography tests, as the subject is one which does not easily lend itself to analysis. Its content is not definite as that of arithmetic or spelling, consequently there is considerable disagreement in the aims and purposes of teaching history. Some would like to have the emphasis placed on names, dates, facts and details; some upon political questions; others upon social and cultural history; and still others upon the philosophy of history, demanding a study of causes and effects of great movements, rather than detailed information; and others upon the study of original sources. For these reasons, the history and geography tests have been developed only recently.

BUCKINGHAM TEST FOR HISTORY, JANUARY, 1916
(University of Illinois)

These tests were constructed for use in the survey of the Gary and Pre-vocational schools of New York City. The values of the questions were determined experimentally. This test is not printed for distribution, but is described in the reference below.

Since then, the author has made a study of the ability to use historical facts. See references below.

Address: B. R. Buckingham, University of Illinois, Urbana, Illinois.

References: Buckingham—"A Survey of the Gary and Pre-vocational Schools of New York City."

Buckingham—"Correlation between Ability to Think and Ability to Remember." *School and Society* 5: April 14, 1917, pp. 443-448.

THE BELL AND McCOLLUM TEST, 1917

(Arranged by L. W. Sackett, the University of Texas)

This test is arranged in a very convenient form in a little booklet. Directions for giving and scoring the test and standards for different types of schools are given on the covers of the folder. The tests included are as follows:

- I. Give the reason for the historical importance of each of the following dates in United States History. (Then follows a column of 10 dates.)
- II. Indicate for what each of the following men was celebrated. (A series of 10 names is given.)
- III. Mention the name of the man prominently connected with each of the following historical events. (10 events are given.)
- IV. Define in short sentences each of the following historical terms. (10 terms follow.)
- V. Make a list of all the political parties that have arisen in the United States since the Revolution, and state the principle advocated by each. (Value 431.)
- VI. Indicate the great divisions or epochs of United States History. (Value 415.)
- VII. On the accompanying outline map of the United States draw the land boundaries of the United States at the close of the Revolution, and indicate by drawing boundaries and naming what have been the different acquisitions of territory since that date. (Value 325.)

The score of each part of questions I to IV is given in the margin. The time allowance for the different tests is as follows:

Test I.....	4 minutes
Test II.....	5 minutes
Test III.....	3 minutes
Test IV.....	7 minutes
Test V.....	5 minutes
Test VI.....	5 minutes

Repeated trials have shown this to be ample time.

Following the usual blanks for name, age, grade, etc., is a space for a personal report on the length of time spent in studying United State history, the time elapsed since the study of that subject, and an expression of the order of preference for English, geography, physiology, history and arithmetic while in school.

Address: L. W. Sackett, the University of Texas, Austin, Texas.

Reference: Bell and McCollum—A Study of the Attainments of Pupils in United States History. Jr. Ed. Ps. 8: May, 1917, 257-274.

THE STARCH AMERICAN HISTORY TEST, SERIES A, SPRING OF 1917 (University of Wisconsin)

This test is devised on exactly the same plan as the Starch Geography Test, that is, upon the principle of mutilated sentences. There are 69 questions in all and most of these contain several blanks to be filled in. The deficiencies of the test pertain chiefly to the nature of questions asked. Many details and isolated facts are called for to the neglect of thought questions, historical reasoning and judgment, and appreciation of the comparative importance of the events. Of course many of the questions are of great importance, however. Standards based on the work of 2,000 pupils are given for grades 6, 7, 8 and high school. The scoring is done by counting the number of blanks correctly filled in.

Address: Dr. Daniel Starch, University of Wisconsin, Madison, Wisconsin.

Reference: Starch, D., and Elliott, E. C.—Reliability of Grading Work in History. *Sch. Rev.* 21: 676-681, December, 1913.

***THE HARLAN TEST OF INFORMATION IN AMERICAN HISTORY**

(C. L. Harlan, Lewiston State Normal School, Lewiston, Idaho)

This test is arranged in 10 exercises, each containing several sub-questions. The test includes the following elements of historical information: names of important men and their connection with great events, historical terms and examples of each, the connection of important events with names of places, thought questions for the practical application of civics, the selection from a list of names of men prominent in certain periods of American history, the dates of important events, events associated with important dates, examples illustrating the truth of four general statements concerning the history of our country, giving the significance of four topics in the history of the United States, and citing important immediate historical results of a list of five causes.

Accompanying the test is a folder of complete directions for using and scoring the test and standards based upon the answers of over 2,000 pupils. These tentative standards for the 7th and 8th grades are for the "end-of-the-year tests." The standards are 56 for the 7th grade and 86 for the 8th grade. A score key is furnished, giving the answer to each question and the value of each element of each exercise. The sum of the values of the questions correctly answered is the pupil's score. The class score is the approximate median as determined by the method described on the accompanying record sheet.

This test has the advantage of including some important elements in history. Some of the questions are purely thought questions and others test the student's knowledge of causes and effects.

Address: Bureau of Co-operative Research, University of Minnesota, Minneapolis, Minnesota.

RAYNOR AMERICAN HISTORY TEST, 1919

(W. H. Raynor)

This test is based upon Bagley's investigation of elementary histories. In plan it is similar to the American History Test by Starch.

Address: Bureau of Educational Research, University of Illinois, Urbana, Illinois.

Reference: Rugg, E. N.—Character and Value of Standardized Tests in History. *School Review*, December, 1919, 757-771.

***THE VAN WAGENEN AMERICAN HISTORY SCALES A AND B, 1919-20**

(University of Minnesota)

This is by far the most extensive and thorough history test yet devised. Its derivation was the subject of a dissertation for the Doctor's Degree. Scale B is an equivalent test for Scale A in case the test is repeated by the same pupils. The series consists of three different scales: an information scale, a thought scale, and a character judgment scale. In addition to these, there is a character judgment scale L which is more extensive and more difficult than scales A and B. The information scale consists of 34 questions, most of these with subheads. The questions are arranged in the order of increasing difficulty. Question No. 17 on scale A is:

"What group of Indian tribes lived in the western part of New York?"

Question No. 17 on scale B is:

"Name two American generals who fought in the Revolutionary War.

"Name one more American general who fought in the Revolutionary War."

The thought scale consists of 22 questions, including many subheads, arranged in the order of increasing difficulty. A typical example of these questions is No. 11, scale A, which follows:

"Previous to the Civil War a large part of the Southern cotton crop was exported to England:

"(a) What was evidently one of the chief occupations of England?

"(b) What effect did the blockade of the Southern ports by the North during the Civil War have upon this occupation?"

No. 11, scale B, is the following question:

"At the beginning of the 19th century voting and office holding in the United States were for the most part restricted to property holders. During the next thirty years, with the growth of manufacturing, the people who worked for wages but owned little or no property became a larger part of the population. These people wanted shorter hours of work and better educational opportunities for their children.

"(a) In order to get these things what would you expect the laboring people to demand?"

The character judging scale consists of 15 questions. Question No. 7 from scale A is:

"In 1724 the Massachusetts Colony determined to put a stop to the Indian ravages. One of their armies of about eighty men under Moulton cautiously advanced through a forest to the open village of Norridgewock. Not an Indian was stirring, till at length a warrior came out from one of the huts, saw the English, gave a startled war-whoop, and ran back for his gun. Then all was dismay and confusion. Squaws and children ran screaming for the river, while the warriors, fifty or sixty in number, came to meet the enemy. Moulton ordered his men to reserve their fire until the Indians had emptied their guns. The savages fired wildly and did little or no harm. The English, still keeping their ranks, returned a volley with deadly effect. The Indians gave one more fire, and then ran for the river. Some tried to wade to the farther side, others swam across, while many jumped into their canoes, but could not use them as they had left the paddles in their huts. Moulton and his men followed close, shooting the fugitives in the water or as they climbed the farther bank.

"(a) Draw a line under the three of the following words which you think best describe the action of the English Colonial soldiers:

frightened	resolute	excited	terrified	careless
deliberate	wavering	timid	cowardly	cool

"(b) Draw a line under the three of the following words which you think best describe the action of the Indian warriors:

treacherous	brave	crafty	excited	cool
terrified	courageous	resolute	bold	irresolute"

Question No. 8, scale B, is:

"The first newspaper published in New York City, the New York Gazette, was the organ of the governor and the aristocratic or court party. Nine years later, in 1734, the Weekly Journal, edited by Zenger, appeared and was from the start the organ of the popular party. At the time the governorship of the colony was being used to pension off any court favorite otherwise unprovided for, without reference to the result of his appointment upon the colony. Zenger began publishing a continuous succession of attacks on the crown officials, the governing class, and finally upon the governor, Crosby, himself.

"Zenger was arrested and thrown into jail on the charge of libel. As the chief justice at the time belonged to the popular party, he was turned out of office and replaced by one of the stoutest upholders of the crown. Even Zenger's lawyers were disbarred from the court, so that he had to be defended by one imported from Philadelphia. The defense was that the statements asserted to be libelous were true. The attorney-general for the crown took the ground that if they were true, the libel was only so much the greater. The judges instructed the jury that this was the law, but the jury acquitted Zenger. The acquittal was hailed with clamorous joy by the mass of the population, and gave an immense impetus to the growth of the spirit of independence.

"(a) Draw a line under the three of the following words which you think best describe the action of Zenger in thus attacking the court party:

spiteful	petty	independent	ignoble	daring
reckless	wavering	foolhardy	patriotic	timid

“(b) Draw a line under the three of the following words which you think best describe the action of the governing class in thus prosecuting Zenger:

brave patriotic unjust courageous prudent
contemptible just judicious despicable careless

“(c) Draw a line under the three of the following words which you think best describe the action of the jury in acquitting Zenger:

unfair just timid traitorous free
despicable submissive cautious independent ignoble”

Character judging scale L contains 10 long questions, including subheads.

A manual by the author gives complete information for giving and scoring the test and standards.

These scales seem to be by far the most promising of the history tests. Their greatest drawback seems to be their expense in time and money, but the advantage gained will probably far more than offset this factor.

Address: Bureau of Publications, Teachers' College, Columbia University, New York.

Van Wagenen, M. J.—Historical Information and Judgment in Pupils in the Elementary School. Bureau of Publications, Teachers' College, Columbia University, New York.

THE GOODMAN AND SACKETT UNITED STATES HISTORY TEST, 1920 (University of Texas)

This series contains eight tests with subheads. These tests are: dates-events, events-dates, names-events, events-names, result-events, causes-events, the selection from a list of events, those occurring between certain dates, and the writing of a paragraph of about one hundred words on one of a list of four topics.

The test is as yet incomplete. Directions for giving and scoring the test and standards are available.

Address: Goodman, H. H., and Sackett, L. W.—University of Texas, Austin, Texas.

THE DAVIS TESTS IN UNITED STATES HISTORY—COLONIAL PERIOD (University of Pittsburgh)

This test, based upon the results of Bagley's study of elementary histories, is still in preparation.

Address: S. B. Davis, University of Pittsburgh, Pittsburgh, Pennsylvania.

THE HAHN HISTORY SCALE FOR GRADES 7 AND 8 (State Normal, Wayne, Nebraska)

This scale is still in preparation.

Address: H. H. Hahn, Wayne State Normal School, Wayne, Nebraska.

THE BUCKINGHAM HISTORY TEST FOR GRADES 7 AND 8 (University of Illinois)

This test is in press at the present time.

Address: B. R. Buckingham, Bureau of Educational Research, University of Illinois, Urbana, Illinois.

GENERAL BIBLIOGRAPHY ON HISTORY TESTS

Foster, H. G.—Adequate Tests in History. *Hist. Teachers' Mag.* 5: April, 1914, 116-123.

Morehouse, F. M.—Testing Results in History Teaching. *Hist. Tch. Mag.* 8: November, 1917, 301-305.

Myers, G. C.—Delayed Recall in History. *Jr. Ed. Ps.* 8: May, 1917, 275-283.

Rugg, E. U.—Character and Value of Standardized Tests in History. *School Rev.* 27: December, 1919, 757-771.

X—PHYSICAL EDUCATION

*RAPEER SCALE FOR MEASURING RESULTS OF PHYSICAL EDUCATION, 1917

This scale is based upon a study of some thirty different types of physical-efficiency measures which the author had collected. Instead of employing a single measure as an index of all desirable changes, the author has devised a five-fold scale or score-card for measuring some of the principal results of physical education, "selected and built up from the many already invented." This scale includes the one devised by Dr. W. L. Foster and reported in the American Physical Education Review for December, 1914; the norms for height, weight and breathing capacity proposed by B. T. Baldwin in the Fifteenth Yearbook of the National Society for the Study of Education, Part I (individual score cards obtainable from B. T. Baldwin, University of Iowa, Iowa City); and Stecher's Physical Ability Scale, in his Educational Gymnastics, J. J. McVey Co., Philadelphia.

The different divisions of the scale are reproduced below.

- I. Health Scale (Rapeer).....total points 25
Count off four points for each serious ailment or defect reported during the school year. For uncleanness reported, count off one.
- II. Physiological Efficiency Scale (Foster's).....total points 15
- III. Physical Development Scale (Baldwin's).....total score 20
- IV. Physical Ability Scale (Stecher's).....total score 25
- V. Mental Qualities Scale (Rapeer's).....total score 15

This test includes ability to co-operate, qualities of leadership displayed, willingness to practice good posture, good hygiene and good, clean living, knowledge of physical education, etc. The points in Scale V are distributed somewhat according to the normal probability curve.

Directions for giving and scoring the tests are given in the reference below.

Address: L. W. Rapeer, 1719 H Street, N. W., Washington, D. C.

Reference: Rapeer, L. W.—Minimal Essentials of Physical Education, and a Scale for Measuring Results of Physical Education. Sixteenth Yearbook of National Society, 1917, Public School Publishing Co., Bloomington, Ill.

XI—SUPERVISORY TESTS

*COURTIS SUPERVISORY TESTS, 1918 (Bureau of Research, Detroit)

These are intended as research tests for teachers and supervisors interested in bringing children of different classes up to the standard for the grade. "They are diagnostic in that they classify the children on the basis of their needs for special attention. Provision is made for keeping a continuous record of each individual throughout his school life. They are really a continuous survey and a means of supervisory control." These tests are generally briefer than the regular research tests by Courtis.

Tests are furnished for arithmetic, composition, geography, writing and spelling. Graph sheets are supplied for plotting the results of the class. The scoring is done in such a way that a class, every member of which is up to standard, will have a score of 1,000 points. Class record cards and individual record cards and class record sheets are furnished for each of the tests and are printed in different colors so that the cards and blanks belonging to a certain test can easily be distinguished.

These tests are different from any other tests available. They fulfill the need of supplying the superintendent with diagnostic material for all of the pupils in school in nearly all of the important elementary school subjects. These are kept on file for reference at any time. The superintendent has in his hand information regarding each class and every individual pupil

in the school system. The tests include an instruction bulletin of 52 pages, and a folder of general information regarding the tests.

These tests are especially useful to superintendents.

Address: S. A. Courtis, 82 Eliot Street, Detroit, Michigan.

Reference: Courtis, S. A.—Standard Supervisory Tests, Folder 1918-19, 82 Eliot Street, Detroit.

XII—CITIZENSHIP

***UPTON AND CHASELL SCALE FOR MEASURING HABITS OF GOOD CITIZENSHIP, 1919**

(Teachers' College, Columbia University)

This test is being prepared as eight separate scales with scoring device. A part of the test consists of a list of over 175 common acts of school children which have been rated on a scale of 1 to 10 according to the combined judgments of more than 70 competent judges.

The preliminary discussion of this scale is published in the Teachers' College Record, January, 1919.

Address: Bureau of Publications, Teachers' College, Columbia University, New York.

Reference: Upton, Mrs., and Chassell, Clara F.—Monograph (new edition being prepared.)

XIII—MUSIC

***THE SEASHORE TESTS OF MUSICAL TALENT, 1919**

(University of Iowa)

This is a very elaborate method of testing musical talent. A complete test requires three days and proper laboratory equipment. An abbreviated form of the test is possible, however, by means of the phonograph. Seashore has put on the market disc records which test five phases of musical ability—pitch, intensity, rhythm, consonants and dissonants and musical memory.

The method may be indicated by a description of the test for pitch. Ten series of ten groups of tones in a series constitute the test for pitch. Each group of the series consists of two tones. The subject decides and records by abbreviation, l or h, on a record blank provided for the purpose whether the second note is lower or higher than the first note. In each successive series the difference in pitch is progressively less up to a certain point, so that it becomes very difficult to decide concerning the pitch of the tone. After that point the difference in pitch is progressively greater, so that the subject is tested twice for each degree of difference.

At the end of the test the correct letters are read off from a score card so that the pupils can check their incorrect responses. This gives the per cent of correct responses on the test. A table is given, by means of which the percentile rank can be determined. This is done for each test separately, and the graph is constructed for the pupil's performance on the complete test. The test may be given to a large number of pupils at once. The author suggests that it should be given at least twice during the elementary school period.

Address: E. E. Seashore, University of Iowa, Iowa City, Iowa. Also: The Columbia Graphophone Co., New York.

References: Seashore—Musical Talent, Macmillan and Co.

Seashore—Musical Talent Chart, University of Iowa, Iowa City, Iowa.

Seashore—Manual of Instructions and Interpretation for Measures of Musical Talent. Columbia Graphophone Co., New York.

GENERAL BIBLIOGRAPHY ON MUSIC TESTS

Baldwin, R.—Efficiency in School Music, Teaching and Practical Tests of Same. Jr. of Proc. Music Supervisors National Conf. 1914, 43-50.

- Bingham, W. V.—Some Psychological Aspects of Public School Music Instruction. Proc. Music Superv. National Conf. 1916: 97-102.
- Gaw, E. A.—Music Tests, Iowa Alumnus 16: May, 1919, 248-250.
- Seashore, in Eighteenth Yearbook National Society, Pt. II, pp. 123, Bloomington, Ill. (Publ. School Publ. Co.)
- Seashore, in University of Iowa Studies in Psychology No. 7, pp. 163, Psych. Mon. 25, No. 2, 1918.

XIV—VOCATIONAL EDUCATION

LEAVITT PRELIMINARY TEST FOR MANUAL ARTS, 1919

This is intended by the author merely as a preliminary test, and is probably not yet in its final form.

Address: F. M. Leavitt, Department of Vocational Education, Pittsburgh, Pennsylvania.

Reference: Leavitt, F. M.—Standardized Measurements in the Field of Industrial Arts. Indus. Arts Mag. 8: April, 1919, 132-138.

WARDNER TEST FOR KNOWLEDGE OF TOOLS, 1919

This test requires the identification of a large number of tools. It is based upon the Army Trade Tests.

Address: C. A. Wardner, Springfield, Vermont.

Reference: Wardner, C. A.—Applying the Army Trade Tests to Vocational Schools. Indus. Arts Mag. 8: October, 1919, 402-403.

XV—RELIGIOUS GROWTH

*THE HARTSHORNE SCALE FOR MEASURING GROWTH IN RELIGION, 1919

(Union Theological Seminary)

The author is of the opinion that "religious capacity is evinced in proportion as a person succeeds in getting his life organized in terms of some valued end or consequence toward which he is working with all his might." In order to find himself in this extended life he must "be able to visualize the end he seeks and to devise means to carry him thither."

A preliminary analysis of factors which are of importance in measuring capacity for and growth in religion is included in the study.

The author's classification of the most important matters to be tested is reproduced here.

- I. The child's practice. The account of this practice in terms of religious functioning.
- II. Factors involved in religious functioning:
 1. Responses to foreseen social consequences.
 - A. Are social consequences foreseen?
 - B. What is the response? Is it socialized?
 2. Self-organization.
 - A. What purposes are formed?
 - B. Are they effective? To what extent do they control conduct?
 3. Ideas and ideals.
 4. Valuations, attitudes, appreciations, motives.

The test suggested for I, The Child's Practice, is the Upton and Chassel Scale for Measuring Habits of Good Citizenship. This scale comprises a list of common acts of children which have been evaluated by more than seventy judges and arranged on a scale of 1 to 10 according to their contribution to democratic school citizenship.

Several tests are suggested for II, Factors involved in Religious Functioning. For 1 the author attempted to find a test which was both an intelligence

test and a test of religious development. For this purpose Mr. John Lacy's test was chosen, an opposite test consisting of two lists of 50 words each, selected from lists of moral traits.

For judging character (2), Scott's plan of judging personnel is used. Character is taken to mean "steady devotion to a cause or social ideal." A score card is prepared by assigning a score of 10 to the child of the teacher's acquaintance who has the best developed character, and a score of 1 to the child who is "least socially developed in character." The children are then rated, boys and girls separately, on a scale of 1 to 10, with these children representing the upper and lower limits of the scale.

For judging ideas and ideals (3), the motives given in answer to the following questions are compared: "What are you going to do when you are grown up? Why?" The answers implying social motives and purposes are accorded the highest rank, those implying selfish purposes, the lowest.

"The ability to discriminate among possible ways of behaving" (4) can be shown by giving the child a situation in the form of a printed story, and a series of cards on which are listed several possible solutions. The child is asked to arrange these in the order of their value.

Another means of testing the same abilities is to arrange from the list of acts in the Upton and Chassell Scale five groups of ten each, ranging in value from 1 to 10 as estimated by the combined judgments of seventy-odd judges. The child is asked to "arrange these things children do in the order of their importance."

This scale is yet in the preliminary stage, but it is already of use, since it contains helpful and definite suggestions for more accurate judgment of religious growth.

Address: Dr. Hugh Hartshorne, Union Theological Seminary, New York City.

Reference: Hartshorne, H.—Measurements of Growth in Religion. Religious Education 14: June, 1919, 148-155.

XVI—SEWING

THE MURDOCH SEWING SCALE, 1919

(Teachers' College, Columbia University)

This scale consists of 15 graded charts, which sell for the price of \$1.00. At the time of this writing it has not yet been possible to procure a copy of the charts and the book from the publishers, and consequently a discussion of the scale is omitted here.

Address: Bureau of Publications, Teachers' College, Columbia University, New York City.

Reference: Murdoch, Katharine—The Measurement of Certain Elements of Hand Sewing. Columbia Contributions to Education, No. 103. Teachers College, Columbia University, New York, 1919, pp. 120.

SOME OF THE PRINCIPAL SCHOOL SURVEYS EMPLOYING STANDARDIZED EDUCATIONAL TESTS AND MEASUREMENTS

Following the name of each survey is a list of the tests used.

Butte School Survey, 1914. Address: Board of School Trustees, Butte, Montana.

1. Ayres Spelling Scale.
2. Hillegas Composition Scale.
3. Curtis Standard Research Tests in Arithmetic, Series B.
4. Stone Reasoning Test in Arithmetic.

The Survey of the Public Schools of Springfield, Illinois, 1914.

Address: Leonard P. Ayres, Russell Sage Foundation, New York.

1. Ayres Spelling Scale.
2. Ayres Handwriting Scale, Three Slant Edition.

3. Stone Reasoning Test in Arithmetic.
4. Ayres Arithmetic Fundamentals (these tests have since been abandoned in favor of the Curtis Tests, Series B, to which they were similar).

The Salt Lake City Survey.

Address: School Survey Committee, Salt Lake City, Utah.

1. Ayres Spelling Scale.
2. Thorndike Handwriting Scale.
3. Curtis Arithmetic Tests, Series B.
4. Hillegas Composition Scale.
5. Kansas Silent Reading Test.

The Denver School Survey, 1916. Address: The School Survey Committee, Denver, Colorado.

1. Willing Composition Scale.
2. Ayres Handwriting Scale.

Cleveland Educational Survey, 1916. Address: Leonard P. Ayres, Russell Sage Foundation, New York City. The volume describing the Educational Tests and Measurements is Judd's Measuring the Work of the Public Schools.

1. Ayres Handwriting Scale, Gettysburg Edition.
2. Ayres Spelling Scale.
3. Cleveland Survey Arithmetic Tests.
4. Gray Test of Silent Reading.
5. Gray Uniform Test in Oral Reading.

School Survey of Grand Rapids, Michigan, 1916. Address: School Board, Grand Rapids, Michigan.

1. Gray Test of Oral Reading.
2. Gray Test of Silent Reading.
3. Willing Composition Scale.
4. Cleveland Survey Arithmetic Tests.
5. Ayres Handwriting Scale.

A Self-Survey of the Sterling Public Schools, 1917. Address: Colorado State Teachers College, Greeley, Colorado.

1. Curtis Arithmetic Tests, Series B.
2. Starch Arithmetic Scale A.
3. Gray Standardized Reading Paragraphs (oral reading).
4. Kansas Silent Reading Test.
5. Ayres Spelling Scale (Curtis Standard Tests in Spelling).
6. Thorndike Writing Scale.

Educational Survey of the Public Schools of Brookline, Mass., 1917. Address: School Committee, Brookline, Mass.

1. Curtis Tests in Arithmetic, Series B.
2. Stone Reasoning Test.
3. Ayres Spelling Scale.
4. Boston Spelling List.
5. Holmes Test for Speed in Handwriting.
6. Holmes Test for Quality of Handwriting.
7. Holmes Test for Speed of Silent Reading.
8. Holmes Test for Quality of Reproduction.
9. Harvard-Newton Composition Scales.

The Idaho Springs Survey, 1918. Address: University of Colorado, Boulder.

1. Monroe Standardized Silent Reading Tests.
2. Gray Standardized Silent Reading Test.
3. Thorndike Reading Scale, Visual Vocabulary.

4. Woody Arithmetic Scale, Series B.
5. Curtis Standard Arithmetic Tests, Series B.
6. Clapp Standard School Tests, Upper Arithmetic.
7. Ayres Measuring Scale for Ability in Spelling.
8. Ayres Measuring Scale for Ability in Handwriting.
9. Clapp Standard School Tests: Correct English.
10. Nassau County Supplement to the Hillegas Scale for the Measuring of Quality in English Composition.

The St. Louis School Survey, 1918. Address: World Book Company, Yonkers, New York.

1. Gray Oral Reading Test.
2. Gray Silent Reading Test.
3. Cleveland Survey Arithmetic Tests.
4. Freeman Handwriting Scale.

Theisen—The Use of Stone Standard Tests in Wisconsin, 1918. Address: W. W. Theisen, Supervisor of Educational Measurements, State of Wisconsin, Madison.

1. Ayres Spelling Scale.
2. Woody Arithmetic Scales, Series A.
3. Thorndike Handwriting Scale.
4. Hillegas Scale for Measuring Quality in English Composition.
5. Trabue Nassau County Supplement to the Hillegas Scale.
6. Kansas Silent Reading Test.

The Educational Survey of Janesville, Wis., 1918. Address: W. W. Theisen, State Department of Public Instruction, Madison, Wis.

1. Woody Arithmetic Scales, Series A.
2. Stone Reasoning Test in Arithmetic.
3. Kansas Silent Reading Tests.
4. Ayres Spelling Scale.
5. Buckingham Spelling Scale.
6. Thorndike Handwriting Scale.
7. The Hillegas, Thorndike and Nassau County Composition Scales.

Report of a Survey of Public Education in Nassau County, New York, 1918. Address: The University of the State of New York, Albany, N. Y.

1. Hillegas Composition Scale.
2. Nassau County Supplement to the Hillegas Scale for the Measurement of Quality in English Composition.
3. Thorndike Scale Alpha for Measuring the Understanding of Sentences.
4. Thorndike Reading Scale A: Visual Vocabulary.
5. Trabue Language Scales C and L.
6. The Woody Arithmetic Scales, Series A.
7. The Curtis Standard Arithmetic Tests, Series B.
8. Stone Test in Arithmetic Reasoning.
9. Thorndike Handwriting Scale.
10. Ayres Spelling Scale.

Survey of the Gary Schools, 1919. Address: General Education Board, 61 Broadway, New York City.

1. Ayres Handwriting Scale, Three Slant Edition.
2. Cleveland Free Choice Writing Test (see Cleveland Survey).
3. Curtis Dictation Test and Composition Test.
4. Ayres Spelling Scale.
5. Misspelled Words in Compositions (Curtis).
6. Curtis Arithmetic Tests, Series B.
7. Cleveland Survey Tests in Arithmetic.
8. Hillegas Composition Scale.