



A Testing Program for Elementary Schools

B. R. Buckingham & Walter S. Monroe

To cite this article: B. R. Buckingham & Walter S. Monroe (1920) A Testing Program for Elementary Schools, *The Journal of Educational Research*, 2:2, 521-532, DOI: [10.1080/00220671.1920.10879084](https://doi.org/10.1080/00220671.1920.10879084)

To link to this article: <http://dx.doi.org/10.1080/00220671.1920.10879084>



Published online: 15 Dec 2014.



Submit your article to this journal [↗](#)



View related articles [↗](#)

Journal of Educational Research

Editorial Office, Urbana, Illinois

Editor-in-chief B. R. Buckingham, University of Illinois

VOLUME II

SEPTEMBER, 1920

NUMBER 2

A TESTING PROGRAM FOR ELEMENTARY SCHOOLS

B. R. BUCKINGHAM AND WALTER S. MONROE

University of Illinois

Most of us are agreed that tests, in order to produce the best results, must be given by teachers. Only in a few school systems can they be administered either by the superintendent in person or by some one acting for him. Simple tests, therefore, will be in greatest demand—tests which are easily administered, objectively rated, and quickly interpreted. Monroe's Standardized Silent Reading Test and Curtis' Standard Research Tests in Arithmetic have been widely used largely because they possess these characteristics of simplicity. The test-maker's art is most clearly revealed when he is able to devise tests which are not only simple but valid as measures of significant school abilities.

The superintendent, therefore, who contemplates initiating a testing program, or even the continuance of one already begun but not carried forward to a point where teachers have become relatively expert, will do well not to be too ambitious in the selection of the number and character of the measuring instruments which he will use. In a certain sense he will have to "sell" the idea to his teachers. He cannot do this if he imposes upon them a confusing and difficult task. It is better for him to select a few tests, each concerned with important activities, than to try to cover a wide range of school subjects. One or two tests carefully administered and competently followed up will satisfy and convince; many tests may do no more than confuse and antagonize.

This would seem to be particularly true in the case of rural schools. When, therefore, the Illinois Association of County Superintendents appointed a "Committee on Standard Tests" and when this committee, pursuant to instructions, met with us to plan a rural-school testing program, we were at once convinced

that such a testing program would have to be simple. Even if general considerations had not dictated such a policy, the relative inexperience and lack of training of rural-school teachers in administering tests would have convinced us that we should have to be content with small beginnings.

At the same time we recognized that the opportunity was unusual. The testing of rural schools on any considerable scale has not been attempted. Their isolation prevents, except at prohibitive cost, the employment of special examiners. Quite generally the teachers have no experience in giving tests; and the long-range supervision of the county superintendent necessarily precludes the careful oversight which alone can surmount these obstacles. More favorable conditions have for the most part confined the test movement to city systems. Yet it is unquestionably true not only that the rural schools ought to be tested, but that they are in a position to profit by the testing to a peculiar degree. The mere fact that the grades are all under one teacher facilitates any regrouping of children which the results of testing may suggest. Moreover, these grades are, or may be, loosely organized. A child may easily—far more easily than in a graded school—be instructed with one grade in one subject and with another grade in another subject. For these reasons and for many others we were eager to meet the county superintendents in their proposal to test the rural schools of the state.

When the committee met with us, we proposed first, that not more than two educational tests should be used; second, that these tests should be in the most important subjects—i.e., reading and arithmetic; third, that if it were in any way possible, the intelligence of the children should be measured; and fourth, that the pupils' scores in the educational tests on the one hand and in the intelligence test on the other should be definitely related to each other. To these suggestions the committee readily agreed.

But what reading test, what arithmetic, and, above all, what intelligence test should we use? As to the reading test there was not much chance for argument. If we were going to have something simple, we should clearly want to use Monroe's Standardized Silent Reading Test or some test of that character. In arithmetic we first considered Courtis' Standard Research Test, Series B; but we afterwards rejected it in favor of two series of

tests, one for grades III, IV, and V, and the other for grades VI, VII, and VIII. This action was taken because of the varied types of examples which such a plan would provide. The selection of an intelligence test constituted the chief problem. Although we now believe that our rural-school program of testing has a wide and general usefulness, it is well to point out that in our relations with the county superintendents our attention was narrowed to the particular project in hand. Since, therefore, the money available for testing in each county was limited, we were obliged to select an inexpensive intelligence test or face the fact that very few superintendents would be able to undertake our program. But no cheap intelligence test seemed to be good enough for our purpose. In fact, we were confronted by the condition—common enough in these days—that the thing which was good enough to be desirable was also too expensive to be had.

The only way out of the difficulty seemed to be the devising of a new intelligence test. Of course, the expense of such a procedure, when combined with that of printing and distributing, might well be greater than the expense of purchasing tests already on the market, but the cost of devising the intelligence test would be a legitimate part of our cooperation with the superintendents. Our final proposal was, therefore, that we would make up a battery of tests in reading, arithmetic, and intelligence, and that we would print all the material in one booklet. The advantages of bringing the material together in this way hardly need be pointed out. By this means a pupil's record on all the tests would be kept together, and the calculation of certain derived measures (to be explained later) would be facilitated.

This battery of tests we agreed to devise, standardize, print, and distribute in quantities sufficient for the use of the rural schools of the state, and to have it ready for them when they opened in September. Moreover, we agreed to prepare a handbook of instructions to teachers, and to have it printed in time for use at a county superintendents' convention to be held at the University of Illinois on the twenty-ninth and thirtieth of July.

This entire project was rendered especially difficult because the school year was already far advanced. The Committee on Standard Tests first met with us on April 2, 1920. The decision to devise and print a testing instrument for use in the rural schools

was not reached until about ten days later; and the time required for printing the tests and for delivering them to superintendents—delayed as delivery was by railroad strikes—carried us well into May. Even with plenty of time at our disposal, the making of a new intelligence test would not have been a task lightly to be undertaken. Such work involves careful investigation of existing material, the formulation of preliminary editions, a try-out on thousands of children, and endless statistical work, typing, mimeographing, and printing. Nevertheless, such a course of action seemed necessary in order to save the situation.

By making a special effort, and with the effective cooperation of a number of Illinois city superintendents, we were able to carry out our plans. Of the intelligence test we used two preliminary editions and of the reading test one. The arithmetic test is a composite of material of known difficulty mostly from Monroe's Diagnostic Test. The final edition of the test booklet is now coming from the press. It consists of 200,000 copies and over 100,000 of them have been spoken for. The teacher's handbook is likewise ready.

In the compilation of the "Illinois Examination," as we have called this group of tests, we do not claim to have displayed any particular originality. Certain materials both for intelligence and educational testing are to be had for the asking, and such material we have used. In this respect we resemble Homer as represented by Kipling in his introduction to the "Barrack-room Ballads":

When 'Omer smote 'is bloomin' lyre,
He'd 'eard men sing by land an' sea;
An' what he thought 'e might require,
'E went an' took—the same as me!

Even if time had permitted us to originate all the material of which the tests are composed, it is doubtful whether our time would thus have been most serviceably employed. At any rate, we have preferred to be useful first and original afterwards.¹

The "Examination" appears in two parts, and is accompanied by record sheets and the handbook of instructions to which we

¹ Dr. Sidney L. Pressey kindly permitted us to use his tests of Verbal Ingenuity and Arithmetical Ingenuity. Professor E. H. Cameron contributed the Substitution Test.

have referred. Part I is for grades III, IV, and V. It comprises the following material:

1. *The Intelligence Test*, consisting of sub-tests in analogies, arithmetical problems, sentence vocabulary, substitution, verbal ingenuity, arithmetical ingenuity, and synonym-antonym.

2. *A Revision of Monroe's Standardized Silent Reading Test for Grades III, IV, and V*.

3. *A Test of Ability in the Operations of Arithmetic*, consisting of eight sub-tests—four on the combinations, and four involving simple examples in each operation.

Part II is for grades VI, VII, and VIII. The intelligence test is the same as in Part I. The reading test is Dr. Monroe's revision of his Standardized Silent Reading Test for grades VI, VII, and VIII. Although this material differs from the reading material for Part I, the instructions are in every way the same. The arithmetic test consists of seven sub-tests as follows: column addition, long multiplication, long division, subtraction, addition and subtraction of fractions, multiplication and division of fractions, division of decimals.

The silent reading test yields measures of both rate and comprehension. For the operations of arithmetic, it was considered desirable, however, to have a single score which would serve as an index of the pupil's achievements. This was secured merely by taking as the pupil's score on each test the number of examples right. These scores were then added after certain adjustments had been made to equalize the weight given to the several tests. The sum thus obtained was the pupil's point score on the operations of arithmetic.

The actual working time of pupils is 27 minutes for Part I and 37½ minutes for Part II. The total time required for giving either of the two parts is about the same—namely, one hour. It is recommended that Part I be given in two separate periods, but that these periods be included in the same half-day session. This is clearly advisable because all the pupils who are present at one test period will then be present at the other.

It is clear that in a rural school (or in a graded school in which children of the fifth and sixth grades meet in the same room) both parts of the examination will have to be administered by the same teacher. This, however, does not mean that two hours of the teacher's time need be devoted to the task. The same preliminary

statement to children and the same intelligence test are used for all grades. Moreover, although the material in the reading tests differs for the two groups of grades, the instructions are identical, and all children may, therefore, be tested simultaneously. Accordingly, the teacher who is giving both parts of the examination may administer them together so far as the intelligence and reading tests are concerned.

The union of intelligence and subject-matter testing has obvious advantages; and these advantages will naturally be more surely obtained if, as in this instance, the intelligence and subject-matter tests are mechanically combined—i.e., form parts of the same testing instrument. But even then the advantages cannot be fully realized unless some easy way is found of combining the scores on the several tests. If the Illinois Examination presented no characteristics other than those we have mentioned above, we should hardly be justified in devoting this article to it. It is, however, in the treatment of scores that the examination is truly distinctive.

The sum of the scores obtained by a pupil on the sub-tests of the intelligence scale yields his "point score" for that scale. On the basis of our returns, we have ascertained for each half-year of age the median point score on the intelligence test. The point score of a pupil may then be interpreted with reference to these medians. For example, it is found that 97 is the median point score of children 14 years of age; accordingly, a child whose score is 97 indicates by his performance a mentality which is typical of that age. Fourteen may then be said to be his "mental age."

This conception of mental age is valuable—much more so than that of chronological age. It indicates the degree of mental development which the child has attained at the time the test is given. Since school work addresses itself primarily to mentality, it is, so to speak, the efficient age with reference to school work. Indeed, it may without much inaccuracy be defined as the present ability to learn. It is, therefore, far more important as a basis of action than the number of years a child has lived.

In childhood, however, mental age like chronological age is constantly changing; but unlike chronological age, its change is not registered merely by the flight of time. For a basis of action which contemplates a continuing educational program, we need

some determination (derived if possible from the point score) which shall be relatively unchanging over a period of years. It has been found that one form of relationship between a child's mental age and his chronological age fulfills this condition. This relationship is expressed by dividing the mental age by the chronological age, or stated another way, by finding the percent that the mental age is of the chronological age. The result is called the "intelligence quotient" (I. Q.). Thus, if a child's mental age as determined from his point score is 9 years and his chronological age is 12 years, his I. Q. is 75.² This expresses the fact that during the 12 years that this child has lived his annual mental development has been but 75 percent of normal. In other words, each elapsed year of his life he has developed mentally but three-quarters of a year. On the other hand, if this child of 12 had a mental age of 15, his I. Q. would be 125. This would indicate an annual mental development amounting to one and one-fourth years. An intelligence quotient of 100 arises from the fact that mental and chronological ages are the same. It shows that mental development is normal. Moreover, since the I. Q. is relatively constant, these percents of development will continue to manifest themselves. This fact affords a basis of prophecy. We may be sure, for example, that certain children cannot advance beyond the sixth or seventh grade; and they should, therefore, have a modified course of study with reference to that fact. On the other hand, we may be reasonably confident that certain children can, if we let them, finish the elementary school while they are still young. It is our duty to plan for them a richer curriculum and an accelerated progress. If their intelligence quotients indicate that they are advancing mentally at a more than normal rate, then they should be advanced in school at a corresponding rate.

In order to make it unnecessary for teachers, when using the Illinois Examination, to compute mental ages and I. Q.'s, we have provided as part of the accessory material a table of which Table I is a section. Its use may be illustrated by supposing that an eleven-year-old child achieves a point score of 57. Note that the entry "57" in the first column is written opposite "10-0" in the second column. This indicates that the pupil in question has a mental age of 10 years. Since the child is 11 years old, note also the column under "Chronological Age" which is headed "11-0."

² Conventionally, the quotient is multiplied by 100 to get rid of the decimal point.

By tracing this column to a position opposite the point score of "57," the entry "91" is found. This is the I. Q. of the child in question.

TABLE I. INTELLIGENCE QUOTIENTS

Point Score	Mental Age	CHRONOLOGICAL AGE				
		10-0	10-6	11-0	11-6	12-0
82	12-6	124	119	113	109	104
77	12-0	120	114	109	104	100
72	11-6	115	109	104	100	96
67	11-0	110	105	100	96	92
62	10-6	105	100	96	91	88
57	10-0	100	95	91	87	84
52	9-6	95	91	87	83	80
47	9-0	90	86	82	79	76

The most distinctive characteristic of educational tests has been the fact that they have been standardized. It has not, however, been generally recognized that our present standards apply only to the median or average scores of a group (usually a grade) and not to the scores of individual pupils. It is true that individual scores may be compared with our present grade standards, but only very limited conclusions may be drawn. For example, the fifth-grade standard for Monroe's Standardized Silent Reading Test for comprehension is 21. The scores of a typical fifth-grade class will range from about 10 to 30. If the median score is 21, we say that the *class as a whole* is up to standard although half of the pupils are above the grade standard and the other half are below it. We do not know whether any particular pupil has done as well as he should or not. A pupil may fail to achieve as high a comprehension score as his intelligence (mental age) indicates he should, and at the same time he may have a score above the grade standard. In other words, a bright pupil may be "above standard for his grade" but below *his own* standard. The opposite also is true. A pupil who has a score "below standard for his grade" may have made a very high score in comparison with his intelligence.

The need for individual standards to use in interpreting the scores of individual pupils is particularly urgent when educational tests are used in rural schools because the number of pupils in any one grade is generally so small that it cannot be representative. Not infrequently there are only two or three pupils. In such cases the median score can have only a very limited signif-

icance. This makes the interpretation of the scores of individual pupils imperative and for doing this individual standards are necessary.

A standard score is the score which a pupil should make on a given test. In the case of silent reading and the operations of arithmetic, two obvious factors, upon which the score which a pupil should make depends, are his intelligence or capacity to learn and the amount of his training or roughly the school grade he has attained. The data obtained by giving both the intelligence test and Monroe's Standardized Silent Reading Test to several thousand children indicated that the amount of instruction which the pupil had received contributed very little to the score which he made on the reading test. It was found that the average silent reading score is approximately the same for pupils of a given mental age whether they are in the third grade or the eighth. For this reason it was deemed appropriate to standardize the educational tests of the Illinois Examination on the basis of mental age. Accordingly, the scores of all pupils of each mental age were assembled and the median taken as standard.³ In this way, the standard for each half-year of mental age was obtained.

A pupil's point score may now be compared with *his own* standard, i.e., the standard for his mental age. However, it will be more convenient, and will also make possible some other things, if the mental age standards are used to translate a pupil's point score into an *achievement age score* (A. A.). This may be done in the same way as point scores on the intelligence scale were translated into mental ages. When a pupil's score on an educational test (silent reading or the operations of arithmetic) is expressed as an achievement age score, his mental age becomes his standard. Since both his score (A. A.) and his standard (M. A.) are expressed in terms of a common unit, they may be compared by dividing in the same way that his mental age was divided by his chronological age to obtain his intelligence quotient (I. Q.). The quotient of A. A. divided by M. A. we have called *achievement quotient* (A. Q.).

Table II illustrates, for a brief range of point scores and mental ages, the device which we have provided for securing these derived measures. The first line of Table II may be read as follows:

³ In doing this it was recognized that the educational tests for grades VI, VII, and VIII were more difficult than the ones for the lower grades, and corrections were made in order to adjust the scores of the more difficult tests to the scale of the easier ones.

A pupil whose rate score in reading is 155, whose comprehension score is 12, and whose arithmetic score is 63 shows on all three of these counts an achievement age of 12 years and 6 months. His achievement quotient for each of the three scores is 124 if his mental age is 10 years, 119 if his mental age is 10 years and 6 months, 113 if his mental age is 11 years, *et cetera*.

TABLE II. ACHIEVEMENT QUOTIENTS

POINT SCORES			ACHIEVEMENT AGE	MENTAL AGE				
Rate	Compre- hension	Arith- metic		10-0	10-6	11-0	11-6	12-0
155	12	63	12-6	124	119	113	109	104
150		60	12-0	120	114	109	104	100
145	11	57	11-6	115	109	104	100	96
140		53	11-0	110	105	100	96	92
135	10	49	10-6	105	100	96	91	88
130		44	10-0	100	95	91	87	84
125	9	40	9-6	95	91	87	83	80
120	8	35	9-0	90	86	82	79	76
114		30	8-6	85	81	78	74	71
107	7	25	8-0	80	77	73	70	67

An achievement quotient of 100 means that the pupil has achieved exactly as well as the average of the pupils of his mental age. If his achievement quotient is 130, it means that he has achieved 30 percent more than the average of the pupils of his mental age; on the other hand, if his achievement quotient is 75, we have evidence that he has achieved only 75 percent as much as the average of pupils of his mental age.

The value of individual standards and the achievement quotient is illustrated in Figure 1. This figure shows the scores for comprehension and the achievement quotients of pupils in a fifth-grade class. The former are plotted along the horizontal axis and the latter along the vertical axis. Each pupil is represented by a dot. Distances from the two axes show the two measures of the pupil's achievement. The grade standard in terms of a point score is indicated by the arrow.

The lowest score in the class is four. If we had only the grade standard, all we could say about this pupil would be that he is below standard and at the foot of his class. His A. Q. shows that in comparison with his own standard he has achieved more than is usually achieved by a pupil of his mentality. In fact

when his mental age is considered, he is one of the "good" pupils in his class.

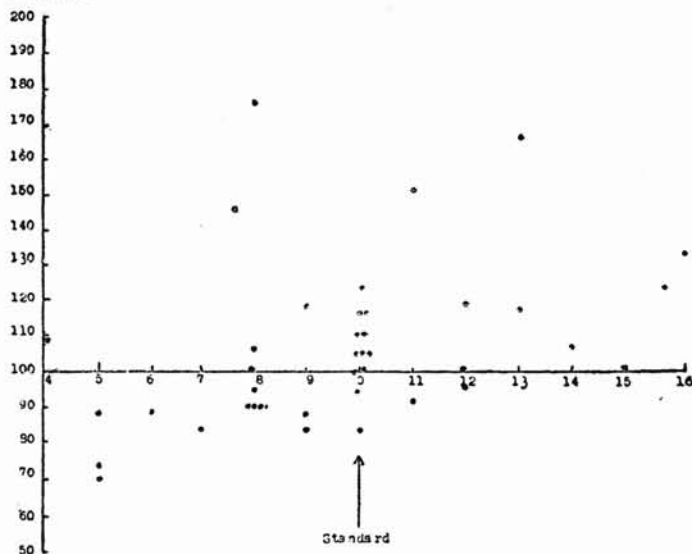


FIGURE 1. SILENT READING COMPREHENSION. RELATION BETWEEN POINT SCORES AND ACHIEVEMENT QUOTIENTS. GRADE V

An added advantage of this plan is that when we have transmuted point scores in subject-matter tests into achievement ages, we have reduced them all to the same units in the sense that each successive year corresponds to an increment of ability gained by typical children in equal lengths of time. Since the achievement ages are expressed in the same units, they may be combined as point scores cannot be combined. For example, we have no obvious way of expressing the total achievement of a child who scores 155 in rate, 11 in comprehension, and 49 in arithmetic. By reference to Table II, however, we observe that these point scores indicate respectively achievement ages of $12\frac{1}{2}$, $11\frac{1}{2}$, and $10\frac{1}{2}$ years. On the assumption that the abilities to which the point scores refer are all equally important, we may obtain the simple mean of these ages, and we may thus express the composite achievement age as $11\frac{1}{2}$ years.

Clearly, the relation of achievement age to mental age is important. If a pupil's achievement age is more than his mental age, it means that some factors other than intelligence are favorable.

The school is doubtless doing its work well. Of course, there are innate characteristics other than intelligence which may be operative. The most important of these is effort; but since any effort greater than usual is not likely to be elicited except through the superior work of the school, it is probably not inappropriate for the school to take to itself credit for the showing.

On the other hand, if achievement age is less than mental age, the same course of reasoning leads us to suspect that the school is at fault. Of course, this may not be true. For example, irregular attendance for reasons over which the school has no control might produce this result. At any rate, when achievement age is less than mental age—i.e., when the achievement quotient is less than 100—an investigation should be made.

Even if the combination of an intelligence test with subject-matter tests were not necessary to secure individual standards, we believe that any competent testing program should involve both kinds of tests. When the latter are given alone and results are compared with standards, great injustice may be done to the school and to the teacher. If the results are below standard, the teacher and the school are likely to be criticized. This may be wholly unjust. The general level of mentality among the children may be below normal by an even greater amount than the scores in subject-matter tests are below standard. On the other hand, scores above standard in educational tests may not indicate that the school is doing good work because intelligence may be above standard to an even greater degree.

Again, the use of intelligence tests without subject-matter tests is one-sided. The real truth about a pupil or a class is made evident when to a knowledge of mentality is added a knowledge of achievement. If mental age may be taken as representing potential ability, achievement age may be taken as representing *effective* ability. The one shows a child's chance to learn, the other what he *has* learned; the one measures his ability as a possession, the other his ability as put to use. Since achievement age indicates ability as it functions, it affords an important supplement to, and commentary upon, mental age as a basis of action—whether that action involves the reclassification of pupils or some other modification of school procedure. Moreover, it affords, when related to mental age, a stimulus to effort or a basis of satisfaction according as it is less or more than the mental age.