

THE PLAN OF MEASURING EDUCATIONAL EFFICIENCY IN BAY CITY

FRANKLIN BOBBITT
University of Chicago

How can the superintendent of a considerable city know whether the results secured in the different ward buildings are kept up to standard in each of them? With so complicated an organization to direct and with such a complexity of results to be secured in every portion of it, how can the superintendent keep his finger on all the work continuously and keep himself sufficiently informed as to the efficacy of every portion of it?

The logic of the present measuring movement in the educational field is simple and irresistible. The problem before the practical superintendent is how to measure the results in all subjects with sufficient economy, expeditiousness, and frequency. Many of the standardized tests are expensive in both time and money. Some of them lose their efficacy if given frequently. For certain of the more important, especially in the more advanced work, standardized tests have not yet been prepared.

The practical superintendent, however, has his daily, weekly, and monthly task before him. He cannot wait until the standardized tests are worked out for the entire field. Accepting the logic of the measuring movement as sound, he must find something practicable to use until the expert measurers are able to give him something better. Even where standardized tests are in the field, he must sometimes employ a rough substitute because of the expensiveness in time and money of the elaborate tests.

An interesting plan of doing the work under such conditions was recently observed by the writer in the schools of Bay City, Michigan. It was devised by Superintendent Frank A. Gause and introduced in the autumn of 1915. After a two years' test, Superintendent Gause expresses satisfaction with the results.

The test, like most of the standard ones, is a modified examination test. In each elementary-school subject a single question is devised as the test of the work of the past month. All of these single questions are then placed on a single examination sheet, and the children write upon all of them as a single test for which a definite and uniform amount of time is given. The following is a sample of the test papers:

GRADE 6 B

Arithmetic:

A farm is 90 rods long and 60 rods wide. Find number of acres in it. Find the cost at \$60.00 per acre.

Grammar:

Fill blanks with correct form:

1. It is. (I or me)
2. We are. (they or them)
3. It was. (he or him)
4. I am. (she or her)

Geography:

Choose one:

1. Make a list of five articles which are manufactured in Bay City.
2. Why did the first settlers come to Bay City?

Hygiene:

Mention some means by which the Board of Health protects the public.

History:

Choose one:

1. What permanent English institution was founded in Edward I's reign?
2. What was the purpose of the Crusades? Briefly tell the story of the Crusader you admire most.

Reading:

Would you like to work for Scoorge? Why?

Spelling:

address	children	thought
promote	doctor	

The question will arise as to the validity of a one-question test. In this connection it must first be mentioned that the test is devised for the measurement of classes and not individual pupils. Superintendent and principals want to find the teachers who are, and

those who are not, keeping their classes as a whole up to standard. Now a single carefully chosen question may not be fair for any particular pupil in a class; but if the question is accurately related to the work of the class as a whole, the law of statistical averages enters in. In a class of forty pupils, even though there is but the one question, there will be forty papers as a measure of the work of that class. The average of the forty, it is claimed, is a fair average of the standing of the class as a whole.

One of the most suggestive features of this plan is the mode of selecting the questions that are used. The city employs a uniform course of study and a uniform series of textbooks. This naturally greatly facilitates the employment of such a testing plan. Whether such uniformity is desirable or not is a question not here discussed. The test is devised with the presumption of reasonable uniformity of curriculum throughout the city. In this respect this test does not differ from the more accurate, standardized tests in general. In practically all cases in their setting up of standards, in their providing for comparison among classes, buildings, and cities, etc., they clearly assume uniformity of curriculum.

But even where there is a uniform course of study and textbooks, because of local conditions there is some, and sometimes considerable, diversity in the amount of ground covered and in the content of the work done. An examination made out in the superintendent's office on the basis only of the prescribed sections of the course of study might be actually suited to the work that is done within one school, and at the same time be a very unfair test for the work going on in other buildings. The plan employed in Bay City took this aspect of the situation into careful consideration. The original questions are not made out in the office of the superintendent. They are made out by the teachers themselves in the various classrooms of the city. Previous to the giving of the test, each teacher makes out a set of questions for each subject which she regards as a fair test of the work that has been done in her class during the previous month. All of these questions are sent in from all of the buildings to the central office. Then from these contributed questions, the special teacher who has charge of the testing work together with the superintendent select for each subject

in each grade one question which, as revealed by the various papers sent in, is considered satisfactory for every building in the city. They are trying, for example, to select a proper question for 6B geography. They take the 6B list of geography questions sent in from the sixteen buildings of the city and select a question that is common to all of the lists, or one which from the questions presented would appear to be satisfactory for each of the buildings. Where it seems impossible to agree upon a single question on the basis of the papers before them, they choose two or three questions of as nearly equal difficulty as possible and permit the teacher to choose the question upon which the pupils will write.

Certain valuable by-products of this method are to be observed. The questions sent in by the teachers present to the superintendent a monthly indication of what the teachers in each classroom consider the essential features of their work. This permits the superintendent to see how well the various buildings are keeping abreast of one another in carrying out the prescribed work. It thus shows him the buildings where overambitious individuals appear to be driving too hard, or doing the work superficially, as revealed by their covering the ground too rapidly. It also shows the laggards in the march and points out where he should bring pressure to bear or improve the conditions by way of speeding up the work. Where buildings are properly moving abreast of one another, the questions reveal the relative emphases in the different buildings. The superintendent thus can know whether it is the essentials that are receiving the major emphasis or the peripheral matters of less significance. When teachers know that their work is currently revealed by the questions submitted as well as by the later results of the tests, naturally they are led to careful thought in the planning of the month's work.

As a matter of fact, one of the objections to the examination plan as a mode of testing results has been that we have never developed any adequate technique for the drawing up of the questions. In Bay City they are devising such a technique. Many of our standardized tests are nothing but modified examinations in which the problems, questions, or other tasks have been selected through the employment of a definite and appropriate technique.

During the first year of the plan, the tests were given every two months. During the past year the tests have been given monthly. We find here a partial answer to the single-question objection. This monthly plan enables each pupil during the year to write upon nine questions in each subject; and a class of forty pupils to present in any single subject a total of 360 papers as the measure of its work in the subject for the year. It would appear that this is a fairly adequate measure of the effectiveness of the work.

The exact date of giving the test is not previously announced to the building. The testing is all done by three specially trained teachers, and no portion of it is given over to the principal or teachers in the building. The latter are present during the test for assisting in the work. The details of the testing are well described in an unpublished report by Miss Adelaide Lawrence, who has had charge of the details of the work since its inception:

Three teachers are given the questions and told to cover the ground as quickly as possible. The city schools, fortunately, are so situated that these three teachers can visit all three buildings in the south end of the city, then move to the central portion of the city with its three buildings, and finally to the north end of the city with its three buildings. This covers all of the buildings on one side of the river. The schools across the river can be covered in like manner in similar sections. There is little chance for the questions to travel on ahead. No principal or teacher in the system is advised as to when the tests are to be given. Upon the arrival of the test teachers in the building, however, word is sent to the various grade teachers as to the testing schedule, and papers are already headed with the name, grade, and school when the tester makes her appearance in the room. Upon her arrival, the test questions are immediately placed upon the blackboard, and from that moment no pupil is allowed to ask a question nor any teacher permitted any comment. A limited time is given for every subject and that time is kept uniform for the grades throughout the city. At the end of the allotted time, the papers, finished or unfinished, are collected.

Upon leaving the classroom they carry the papers with them. All of them are taken to the central office, where they are marked by these special teachers themselves. This permits a uniformity in the marking standards employed for all of the buildings in the city.

The results are tabulated in such a way as to show the relative standing of the different classes throughout the city. These are

combined so as then to show the relative standing of the different buildings. The superintendent thus gets a line on the relative efficacy both of the teachers and of the principals.

As these comparative statements are drawn up, it cannot be expected that all buildings of the city, differently situated as they are, shall all rise to the same level of accomplishment. Naturally they will distribute themselves from high to relatively low. As the monthly summaries are drawn up, however, month after month and year after year, it will be possible to discover the relative level that seems appropriate for each particular building. The monthly comparative results can then be used for the purpose, not of seeing whether all of the buildings of the city are rising to the same level of accomplishment, but whether in the general distribution which seems appropriate for the city each building is holding its own, or in other words, keeping the level which has been shown to be its appropriate position.

Three types of reports are sent monthly to the buildings. The first type is shown in Table I.

To each building is sent the record of its accomplishment in the test. In Table I the heavy-face figures present the record of the Trombley School—one of the schools that in this test stood very near the top of the list. The light-face figures in the table show the average for the entire city in the same test. The principal upon receiving such a report can read across the horizontal lines for each of the grades and see whether or not the work in his building is keeping up to the average. By reading across the lowest horizontal column he can discern whether he has kept the rank which previous experience has shown to be appropriate for his school. This table shows that the Trombley School in this particular test stood sixth in arithmetic among the sixteen schools of the city; fifth in language; first in spelling; second in geography; second in hygiene; first in history; first in reading; and second in its general ranking in all subjects. The principal can read down the vertical columns and discover the grades that appear to be doing the best work in his building as compared with the average of the city, and those that appear to be falling below the standard.

It is these records that in the ultimate analysis will determine whether the one-question plan of giving the test, or whether any such examination test, even if more questions are used, will actually serve the purposes for which this work is done. There is really considerable inertia in the standing of any given class throughout the year or of any given building. If these reports over a series

TABLE I
AVERAGES OF THE TROMBLEY SCHOOL

Grade	Arithmetic	Language	Spelling	Geography	Hygiene	History	Reading	Average
5B	73 100	93 98	88 94	87 96	81 99	85 97	95 100	86.11 97.71
5A	82 93	94 98	92 100	88 90	59 97	91 99	91 97	85.38 96.28
6B	44 90	93 98	98 100	90 91	73 77	91 99	94 99	83.29 93.42
6A	50 0	89 94	88 94	85 90	91 83	90 98	92 99	83.54 79.71
7B	87 100	84 88	95 96	82 88	81 91	90 99	89 99	86.88 94.42
7A	93 100	92 96	94 100	90 98	92 95	93 99	99 100	93.36 98.28
8B	95 100	87 87	95 98	85 84	92 99	97 100	92.11 94.66
8A	93 100	91 88	94 100	84 70	94 99	92 99	91.47 92.66
Average	77 85	90 93	93 98	87 92	81 87	91 99	94 99	86.76 93.39
Rank	6	5	1	2	2	1	1	2

(Averages of the Trombley School in heavy-faced type. City averages in light-face type. Decimals in most cases omitted in the printing.)

of months and years show that the relative standing of certain buildings should be high, others medium, others low; and if the tables then reveal month after month a relative inertia of classes and buildings, the validity of the plan will thereby be indicated. In the degree in which buildings fluctuate from month to month in their relative standing, being now high and now low, and then

again high, the evidence is against the validity of the method. Figures at hand do not yet present conclusive evidence one way or the other.

After each monthly test a list of buildings like that of Table II is sent to all of the buildings. It represents the roll of honor. It shows the buildings that have ranked highest in the work of each grade in each subject and in the averages of grades and subjects. Reading across the horizontal lines each principal can know where the best work is presumably being done in each grade and subject. The lowest horizontal column shows that in the average in all grades in arithmetic the highest ranking building is the Riegel; in language, the Corbin; in spelling, the Trombley; etc. Where the two averages coincide, we find the names of the buildings that have stood highest in the total results of the test. In this particular test the Riegel stood first among the buildings of the city; the Trombley, second; and the Farragut, third. The names of those receiving mediocre or low ranking are not sent out.

One of the values of this report mentioned by the superintendent is the stimulation of the various buildings. Each likes to see its name upon this roll of honor as numerous and frequently as possible. A second value relates to the bringing up of the work in buildings where it is shown to be poor. One of the best modes of training the weak teacher is to have her go and observe work of the character that she ought to be carrying on in her own classroom. Such a list shows where the visiting teachers can be sent with the greatest promise. Naturally this argument is somewhat vitiated by the different standards appropriate to schools in different situations. The very best teachers may sometimes be in a school that never can hope to reach the top in the character of the results secured. This points to the possible desirability of making some modification of the plan of determining the monthly roll of honor.

The third type of report sent to each building shows the names of pupils who have failed upon the examination. No claim is made that this presents an adequate test of individual pupils. It is some help, however, especially after the records have accumulated for a number of months. Each teacher presumably knows who the poorest pupils of the class are. But when the class record has been

TABLE II
THE HIGHEST-RANKING BUILDINGS

	Arithmetic	Language	Spelling	Geography	Hygiene	History	Reading	Highest Average
5B	Fremont Corbin Trombley	Riegel	Riegel	Sherman	Corbin	Lincoln Trombley	Farragut Kolb Wenona	Trombley
5A	Woodside Washington Corbin	Garfield Park Corbin	Sherman Washington Whittier	Trombley	Trombley	Trombley	Trombley	Trombley
6B	Riegel	Riegel Sherman Dolsen	Riegel Park Corbin	Fremont	Trombley	Trombley	Wenona Park Trombley	Riegel
6A	Farragut Garfield	Sherman Washington Whittier	Sherman	Fremont	McKinley	Fremont Trombley	Kolb Trombley	Farragut
7B	Dolsen Whittier McKinley	Whittier	Washington McKinley Whittier	Riegel	Farragut	Fremont Trombley	Sherman Garfield Trombley	Trombley
7A		Kolb	McKinley Trombley	Sherman Fremont	Trombley	Fremont Trombley	Sherman Farragut Garfield	Trombley
8B		Riegel	Sherman Park		Sherman	Lincoln Kolb Wenona	Woodside Sherman Dolsen	Farragut
8A		Garfield McKinley	Sherman Riegel Trombley		Woodside Riegel	Woodside Farragut Wenona	Trombley	Riegel
Highest Average	Riegel	Corbin	Trombley	Fremont	Riegel	Trombley	Trombley	Riegel Trombley Farragut

TABLE III
 FAILED IN TROMBLEY SCHOOL

Arithmetic	Language	Spelling	Hygiene	Geography	History	Reading
5A S. Gonyaw—o 6B	6B C. Bosse—75 7B	5B J. Bowden—6o 6A	6B C. Bosse—5o M. Alexander—7o	5B L. Evans—7o A. Larson—5o	None failed	None failed
B. Frazier—o 6A All failed	8B R. Mattison—6o F. Miller—7o C. Hitchcock—7o 8A T. Doyle—7o	E. Headings—6o	F. Doyle—7o 6A E. Headings—7o 8A J. Merry—7o T. Doyle—65 A. Rosebush—7o			

brought down by low averages on the part of individual pupils, she wants to know exactly who has been responsible for bringing down this class average. The sending of this list of names to her enables her to focus her efforts upon the points which have been actually responsible for the losses; and to stimulate her in her efforts to bring up others in the class who are likely to do the same sort of thing upon the next test.

The superintendent keeps two types of graphic charts in his office for his own assistance and for the inspection of all who care to examine them. Chart I shows the relative standing of each building in the city for the six tests given during the first semester of the year 1916-17. Each number represents the name of a building. The chart shows the average level of standing in all subjects tested. The September standing was relatively low, and the buildings were much scattered along the scale. Either the test was unusually difficult or the pupils at the end of the first month had not got under way. But since for supervisory purposes relative standing is rather more important than absolute standing, this may not greatly matter. If the test is a valid one, such a chart ought in time to reveal the rightful relative level of each building; and whether the buildings each month come up to their several expected relative levels.

Another type of large graphic chart kept permanently in the superintendent's office is illustrated in Chart II. This shows the relative standing of the arithmetic teachers in the city system as regards their ability to get results. It permits any teacher of arithmetic, or rather any group of arithmetic teachers in a single building, to see how they stand as compared with the other arithmetic teachers in the other buildings. The chart shows a very considerable degree of consistency in the relative standing of teachers in different buildings throughout the semester. The work of the two teachers in Building 8 is consistently low. That of the two teachers in Building 15 or of the three teachers in Building 10 is consistently high. Except for the fourth month, Building 6 is consistently mediocre. Although the chart shows Building 4 apparently to plunge about rather violently, yet, as a matter of fact, the record is consistent, except for the single month of test

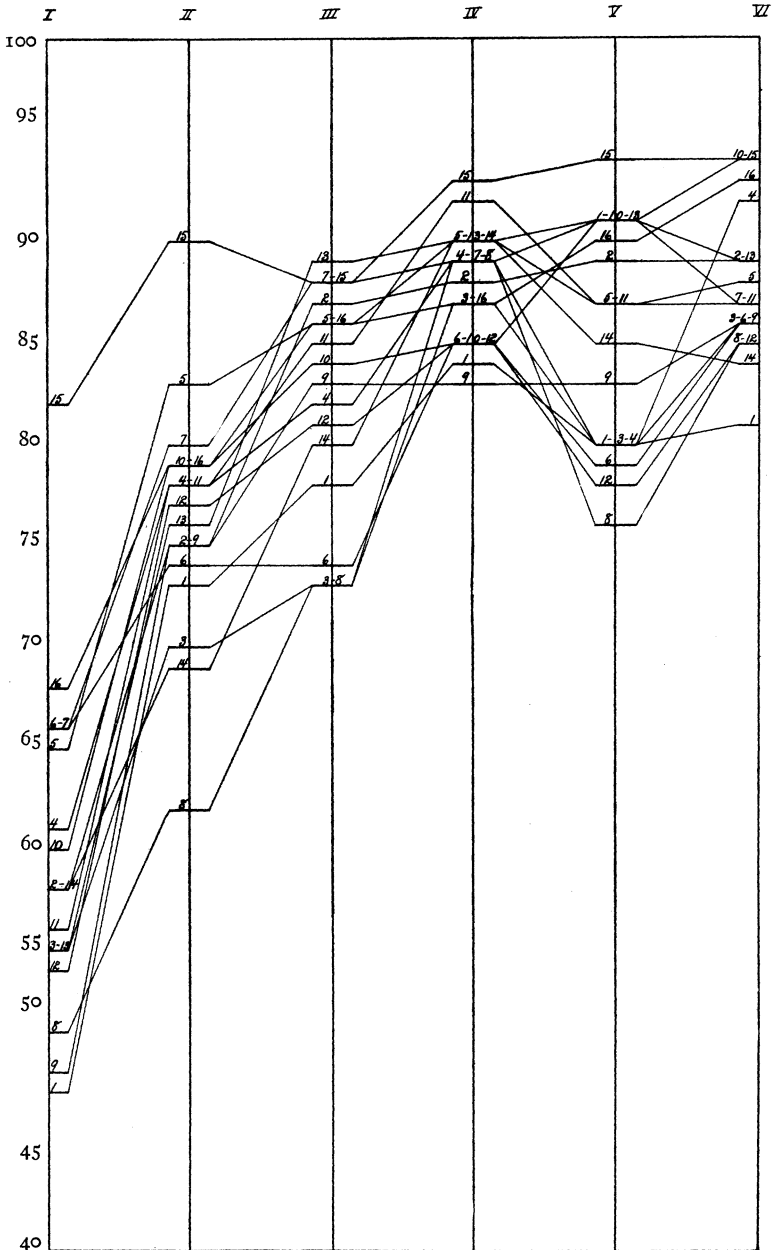


CHART I.—Relative average standing of the sixteen buildings in six consecutive monthly tests.

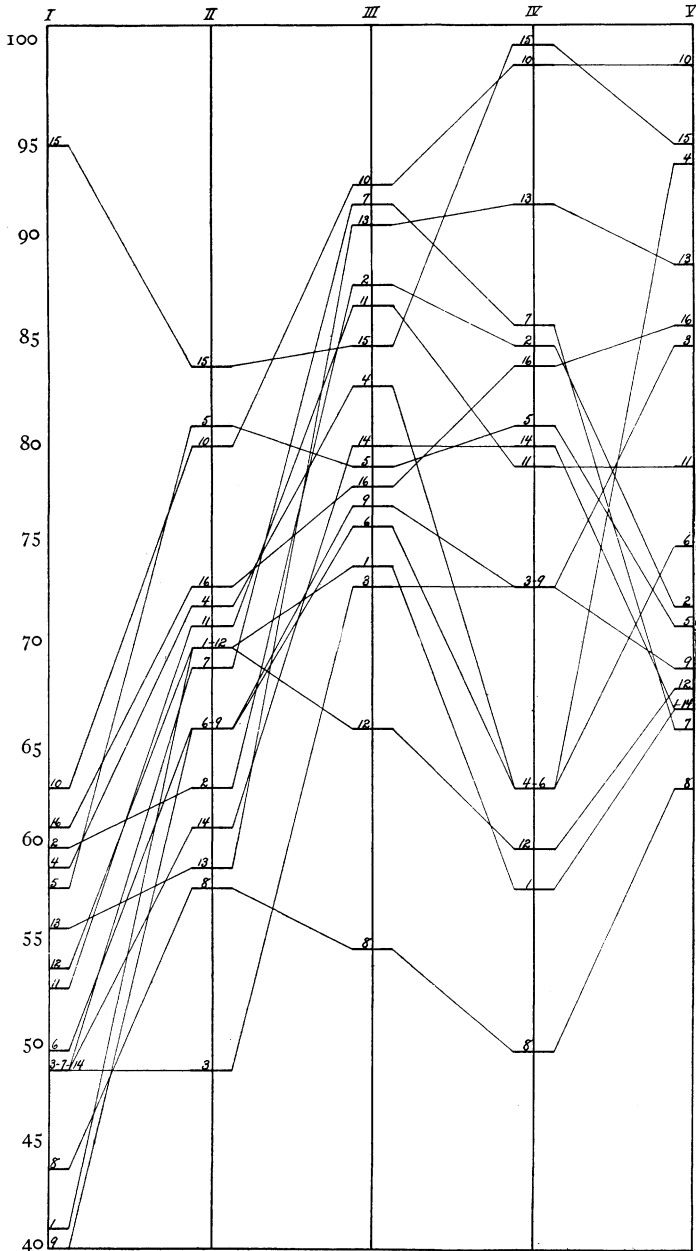


CHART II.—Relative standing of the sixteen buildings in *Arithmetic* in five consecutive monthly tests.

Number IV. It appears that some untoward circumstance entered in to give a probably wrong standing for that month. The perfecting of the plan naturally must lead to the elimination of most of such possibilities of accidental standing. Other charts, similar to this for arithmetic, are kept for each of the other subjects.

Superintendent Gause refers to this as a "tentative plan." He urges no finality for the details. It is clear, however, that it possesses many excellent features that must inhere in any well-developed system of measurement. Like the educational measurers of every type, one of his major tasks now is the discovery and elimination of shortcomings and defects. The various programs of measurement attacking the problem from various angles will tend to complete each other and in the end to bring forth a finished workable plan.