

C
O
P
Y

Princeton University
Princeton, New Jersey
January 3, 1938

President James B. Conant
Harvard University
Cambridge, Massachusetts

My dear President Conant:

A few days before Christmas, Dr. W. S. Learned invited me to lunch in order to tell me of the plans which he is drafting for erecting a mammoth organization to absorb all existing testing agencies. I registered as emphatic a protest as I knew how against his plan, but apparently I made no impression. Since you have gone on record as favoring some sort of consolidation, and since your name will undoubtedly be linked with any plan which may be promoted, I wish to lodge my protest with you, and send a copy to Dr. Learned.

One of my complaints against the proposed organization is that although the word research will be mentioned many times in its charter, the very creation of powerful machinery to do more widely those things that are now being done badly will stifle research, discourage new developments, and establish existing methods, and even existing tests, as the correct ones.

The history of the testing movement shows that the wholesale extension of practical applications has in virtually every instance retarded progress by discouraging further investigations. It seems inconceivable that such a situation should hold, but it results from the fact that the consuming public, as a result of an effective program of propaganda by the test peddlers, believes that the usefulness of a test depends upon its wide standardization. But a test once standardized is fixed permanently and cannot be changed or improved.

The early work of Binet in intelligence testing illustrates the dangers of premature standardization. If the separate questions of the Stanford-Binet test, the most widely used American version, had been studied and analyzed further and revised on the basis of the study, there would probably not be one single item left to be given and scored just as it was when the test was published in 1916. But the test was standardized, it was sacred, and it had to be left entirely alone. Although it must have been given over a hundred thousand times and each administration could have contributed systematically to the improvement and correction of the instrument, the system permitted no change. A revised version of the test has finally, after over twenty years, been separately developed, but whether or not it is any better than the original remains to be seen.

January 3, 1938

Just before the war, Arthur S. Otis had been fooling with group tests, and the army testers used his materials, with others contributed or hastily concocted, with the resultant Army Alpha. This test is an atrocious one, but it was given over a million times during the war, and it is still used in its original form because so many norms are available. I fail to see any significant or fundamental value in these norms, but they have a great sales appeal.

Shortly after 1925 Mr. and Mrs. Thurstone built, or received from contributors, the early forms of the American Council Psychological Examination. This test is given annually to thousands of college students and is given practically unchanged in form from year to year. I am sure that Mr. Thurstone feels that the test is obsolete and should be killed, yet in the face of the popular demand which has been created, he apparently can do nothing. The American Council must realize that the test is obsolete, but it is one of their projects which is "successful" since the test is widely used and they pick up a little income from it. They continue to publish it year after year. Here you have the combination of one of the ablest research minds in the testing field (Thurstone) and a supposedly responsible educational foundation caught in a web of their own spinning from which they are unable or unwilling to extricate themselves.

A review of the testing movement convinces me that large-scale promotion of faulty measuring devices will make it more difficult to develop good ones and will effectively block research. The test market demands something which is cheap and which takes a minimum amount of time to give and score. The first products of research are usually clumsy and bungly. It takes years to iron out procedures and even then it may be impossible to do some things quickly and cheaply. And clumsy expensive procedures necessary for gathering information cannot compete with cheap and sloppy methods which for certain crude group comparisons give what are advertised as "reliable" results. The history of testing proves this. The record of the Carnegie Foundation supports the point also, for they allowed only thirty minutes for what they called "a measurement of intelligence" in their Pennsylvania Study.

Dr. Learned and I long ago agreed to disagree on the results of this Pennsylvania Study. He argued in favor of his conclusions; I argued against his tests and the methods used in drawing conclusions from his data. He finally agreed that his conclusions could not be justified from his data, yet he insisted that the conclusions were good propaganda for the educational world and he has continued to preach them. We agreed to part company amicably on the difference between the point of view of the scientist and that of the propagandist. On the present issue of extending the bounds and influence of faulty and limited measuring techniques, Dr. Learned feels that certain desirable social results may be obtained even though the methods of obtaining them are wrong. In my opinion no social advantages will in the long run accrue from doing anything wrong.

Even in the Foundation's latest venture in examining, their examination at the graduate level, they have scrapped every suggestion for measuring an individual's ability in his field and have been content to test for miscellaneous disabilities, insofar as they are revealed by very brief informational tests supposedly covering a large part of the range of human knowledge. Every time a program of measurement comes up for adoption the longer more painstaking approach is scuttled in favor of the shorter and easier.

In a recent plea for research, a copy of which I am enclosing, I have attempted to suggest means of avoiding some of the dangers inherent in any consolidation into a powerful organization. As an upshot of this article, W. V. Bingham has written me:

What have you decided to do about making available for general use in secondary schools one or more forms of the Scholastic Aptitude Tests? You have developed the best examination there is for measuring academic aptitude. That the measurements which such an instrument yields are not readily accessible to high school principals and other advisers when counseling students about their educational plans is in my opinion one of the scandals of the present situation in the testing movement. I wonder whether the College Entrance Examination Board realizes that in not releasing such examinations it has in part been responsible for the continued wide use of inferior instruments?

To reply to a charge of this sort, it is necessary to review the procedures which were used and which are continually being used in developing the Scholastic Aptitude Test. Our first form in 1926 contained nine sub-tests on which the candidates were rated and sub-test ten which was not scored until later. There were seven different sub-tests ten, each containing a different type of test material for investigation and possible future use. In 1927 we again used sub-test ten for experimental material and this year we had eleven sections in circulation. In 1928 and 1929 we made drastic changes in our methods but still carried experimental sections. We thus had experimental tentacles out feeling our way as we went.

While we were developing the verbal Scholastic Aptitude Test, we were also experimenting with certain devices for describing elementary forms of mathematical insight. The results seemed promising and in a moment of great weakness I reduced the time of the verbal test, introduced the mathematical test, and threw out all experimental sections. The verbal test took seventy-five minutes, the mathematical test seventy-five or eighty minutes, and no time remained for the experimental sections.

This two-axis test was a great success, but it robbed us of every chance of developing either test more thoroughly. And we were forced to continue this practice until 1936 when we restored the longer verbal test with its

January 3, 1938

experimental sections and turned all of the mathematical material over to the four-hour alpha beta gamma mathematics attainment test. We now have English people helping us with the experimental material for the Scholastic Aptitude Test and mathematicians actually developing the mathematics test. Our soul is saved, but we are spending twice as much time and at least twice as much money in doing a job which we did just about as well, for purposes of consumption, as we did from 1930 to 1935. The college deans and our test users were perhaps as happy then as they are now, perhaps happier. They didn't know that we had cut our moorings, and probably didn't care. Having once adopted a short cut, I had to fight my own Board for four years to get back on the harder path.

Now why is the harder road the wiser one? And this is the point that people cannot or will not see. Because society is changing, the language changes, thinking changes, education changes. If one locks a subject as it is taught today with factual tests derived from the teaching of today, one will apply a terrific brake to the advancement of thinking in that subject. It is possible to devise tests empirically from any given academic generation, but no particular generation must be allowed to cast a blight on future generations. One should go the costly way, the long way, the time-consuming way, to build for the future.

The answer to Bingham is that although we could do a good job today with a short form it would not work tomorrow, or if it did work tomorrow we should have succeeded in applying a brake to educational progress. As further evidence of this point, I have recently been told that some of the available standardized chemistry tests which were made less than ten years ago have recently been found to contain, in over one-fourth of their items, material which is now classed by chemists as out of date. The tests are still being used.

The real threat of Dr. Learned's proposal to extend the bounds and influence of testing of a certain type is a threat to future generations. At the present time, the College Board is perhaps the most influential testing organization although in terms of numbers it examines comparatively few students. The Board is now painfully outgrowing its earlier sins and working toward a position in which it does not impose its pattern on education, in which, in fact, it continually derives its pattern from education. It is committed to no particular form of tests and has no conclusions which it is attempting to establish. It has never resorted to pressure methods of selling its services.

The tests of the Board have been pushed into places of prominence as goals of teaching in a way which most of the more limited tests of the newer variety have never been. Fortunately these newer tests are usually given incidentally somewhere along the school course. Teachers, thus far, have not trained pupils specifically for them. If the unhappy day ever comes when teachers point their students toward these newer examinations, and the present weak and restricted procedures get a grip on education, then we may look for the inevitable distortion of education in terms of tests. And that means that mathematics will continue to be completely

President James B. Conant--5

January 3, 1938

departmentalized and broken into disintegrated bits, that the sciences will become highly verbalized and that computation, manipulation and thinking in terms other than verbal will be minimized, that languages will be taught for linguistic skills only without reference to literary values, that English will be taught for reading alone, and that practice and drill in the writing of English will disappear.

Any person or organization which assists in the exploitation of existing procedures and the simultaneous throttling of research and developmental activity must be ready to accept a large responsibility.

Yours sincerely,

Carl C. Brigham