

## COMMUNICATIONS AND DISCUSSIONS

---

### FURTHER SUGGESTIONS REGARDING MENTALITY TESTS

I have tested several thousand children during the past two years, employing the Binet test and many other mental and educational tests, and I wish to record some of the convictions that have grown out of my experience. In the first place, I agree quite heartily with Professor Yerkes that the age arrangement or method of scoring is not so desirable for scientific use as a method of scoring which gives cumulative credit for each additional element of the test accomplished. Giving an additional credit for each successive task performed will tend, however, to distort the facts as to the relative achievements of two individuals unless these tasks are so selected as to fall at practically equal intervals along a scientific scale for difficulty of performance. If each X in the line shown below represents a task to be performed, and its position in the line represents its location upon an accurate scale for difficulty, the facts regarding relative performances will not be greatly distorted by giving one credit for each task performed.

X		X		X		X		X
Easy								Hard

If the same method of scoring were used in the case of the tasks represented below, however, great misunderstanding would result.

X		X		XXX		X
Easy						Hard

To say that Individual A had performed three tasks, while Individual B had performed five, would give an incorrect idea of the relative achievements of A and B in this latter instance, although such a statement would be fairly accurate if the tasks were separated as to difficulty by equal intervals, as shown in the first illustration. We must therefore take care, in arranging tasks in series, to know that the differences in difficulty between consecutive tests are relatively equal.

The problem now resolves itself into one of measuring accurately the differences in difficulty between various elements of a test series. I know of no other satisfactory way to determine the relative difficulty

of two such elements of a test than to try both upon a large number of appropriate individuals. The amount of difference needs, however, to be expressed in such terms as will apply equally well to the difference in difficulty between any other two elements of the test, and in such terms that any other investigator may verify the amount by trying the two elements upon a group of individuals chosen in the same manner as the group originally tested.

In certain educational tests we have assumed that ability to accomplish tasks was distributed within a given school grade according to the "normal surface of frequency," and that one such grade was just as variable as any other, because selected in the same manner by equally competent teachers. Some measure of the variability of a school grade group has been used, therefore, as a measure of differences in the difficulty of elements in the test series. It seems probable that some such function of the normal surface of frequency will be used as the common measure of difficulty for tests applicable to school children and young people. For infants, however, we shall need to find some other measure of differences.

The strongest conviction that has come to me in my work with tests is that we must have tests to measure particular abilities, rather than tests of general mental ability. Instead of reporting a child as mentally defective, we must point out the respects in which he is deficient, and to what degree. A physician does not diagnose a case as merely "a serious illness," but he seeks to discover the respects in which the patient is ill, and to base his treatment upon these particular symptoms rather than upon "poor health in general." Psychologists and educators must likewise be able to point out in what respects a boy will probably be defective for life, and must then seek to adjust training to the individual's possibilities.

As an illustration of the above statements, I found fifty boys in certain institutions of New York State who by the Binet tests were clearly three years or more below normal children of their age. A language completion test applied to these same fifty boys showed that forty-eight of them were three years or more below the normal child of the same age in their ability to complete incomplete sentences. It would have been unfair, however, to have condemned these boys as totally unpromising, for a test of mechanical construction ability applied to them showed that thirteen of the fifty were able to accomplish as much or more than the average child of the same age in the New York City public schools. It is true, of course, that good qualities go

together, and that ability to do mechanical construction work is probably closely related to ability to succeed in the Binet tests; but what we need, it seems to me, is a large variety of tests along many different lines, each test being carefully graded to show how difficult a task of that particular kind the child can accomplish.

M. R. TRABUE.

Teachers College, Columbia University.

---

Within the past few years one phase of my official duties has involved the applying of mental tests to several hundred subjects ranging from kindergarten children to college students. About 150 of these subjects were tested individually by me, over 100 cases involving my applying the Binet-Simon scale. Within this time also, the training and constant supervision of some 120 Binet examiners has devolved upon me. The following opinions are the result of this experience.

The criticism of the Binet-Simon test methods appears to have been due in part to the following causes: (1) The failure to distinguish clearly between the ability to administer the tests scientifically and the ability to interpret the results of such tests in the light of established or partly established norms, clinical data, etc. Graduate students have been turned loose with a syllabus-manual as practically certified Binet testers. They have registered so-called mental ages which educators, jurists, and others have been left to accept as true mental ages without regard to racial or social norms.

(2) The over-idealism of those who have been hoping (a) to establish a universal scale of mentality tests such as Yerkes heralded last year but has since abandoned, or (b) a national American scale such as Terman has been arranging.

(3) Also another form of over-idealism which prohibits interest in a provisional uniform method because statistical data are lacking as to the correctness of each detail of procedure.

(4) The absence of a laboratory manual of the Binet-Simon tests which attempts to answer provisionally the hundred and one questions of procedure in matters of test order, second chances, uniform scoring, etc., involved in every Binet examination. Yerkes' manual is an advance in this direction for those who are willing to abandon the mental level principle.