THE JOURNAL OF EDUCATIONAL PSYCHOLOGY

A CONSTRUCTIVE ABILITY TEST

TRUMAN LEE KELLEY

University of Texas.

Contributions from the University of Texas Educational Laboratory, No. 8.

The number of tests which have been devised to measure general and specific traits of intelligence is so great that any new test which calls for a hearing may properly be expected to measure a new trait, or an old one more thoroughly. Before making such a claim for the present offering, a brief survey of somewhat similar tests is in order.

Many tests have been drawn up specifically with a view to securing a procedure which would not be dependent upon language, nor affected by preceding training. A test cannot be entirely free from environmental influences, but several have been contrived which are independent of preceding school training and familiarity with English or any other particular language. Manipulative and constructive ability tests in the main meet these conditions. The Sequin-Goddard form board, Fernald-Healy form boards, Healy puzzle box, Healy picture completion test, Thorndike-Stenquist constructive ability test, and many of the tests used at Ellis Island, for example the Knox cube test, are of this nature.

All of these tests are alike in that they measure the capacity of the child to accomplish an assigned task. In every case the end is set by the experimenter or by the test. As tests, the object of which is to solve a given problem, they are excellent, but they do not measure the ability to initiate a problem.

The constructive ability test here presented attempts to measure the ability to initiate as well as to execute a task. With such an object in view, the end cannot be set for the child. Another requirement of a useful test is that it shall be capable of accurate grading and standardization. It has therefore been attempted to devise a test which (1) is free from dependence upon language, (2) is free from preceding formal training, (3) tests initiative as well as manipulative ability and (4) is capable of objective grading and standardization.

It may seem that the condition permitting objective grading and the one allowing for individual initiative are mutually exclusive. This has been the point of greatest difficulty in devising the present test and the means of grading it. The test and procedure finally evolved aim to meet both conditions.

The aim has been to have material that would test a wide range of material ages; that would give the minimum of familiarity, yet ample opportunity for building; and, finally, that would be so limited in amount that the time necessary to give the test would not be prohibitive. Each of these points was the subject of careful preliminary investigation, and the resulting was the selection of the test material shown in Plate I, and the adoption, as a time limit, of twice as many minutes as the child is years old. (In case of exceptional children twice the mental age is allowed.) In case the structure is completed in less time, and the child shows no interest in further building, he may be allowed to stop.

In giving the test the material is placed before the child, who is instructed in his native tongue, "to build the best thing he can and whatever he chooses." A number of children tire of building after a few minutes and accomplish little thereafter. In such cases the best attempt is the first or second. Usually the merit of the second or third structure started is the highest and in very few cases would added time for building result in securing structures of greater merit. Ten minutes for a five-year old is ample time, forty minutes for an adult or twenty-year old may be a trifle short, but by that time enough has been built to judge of the probable merit of the structure which would have resulted with slightly more time. With boys fourteen years of age the best structure built has usually been completed by the expiration of twenty-two minutes, so that twenty-eight minutes seems an

PLATE I. TEST MATERIAL

ample maximum time. Structures that are incomplete because of lack of material are not to be graded as high as structures incomplete because of lack of time. The former incompleteness is due to poor judgment in not recognizing the limitations of the material, while the latter may be due to the complexity of the structure conceived. In general the latter situation will arise only when an elaborate structure is started near the end of the allotted time, for the material is too limited to permit of many structures of such complexity that they cannot be built in the time allowed even to a ten-year old child.

The record blank for the test is shown in Plate II. Under "History" is to be entered notes obtained from the child as to blocks and building outfits that he has owned or played with extensively in the past and as to vocational or manual training experience that might be expected to have elements that would transfer to the test situation. The time of starting the test is recorded immediately following the word "time." As soon as the first structure started is completed or abandoned the time is

PLATE II.

Name..... Age..... Date of birth..... Date.....
 History..... Grade.....

Time							Final time
Purpose	1.	2.	3.	4.	5.		
Dominance of purpose							
Symmetry							
Interest in building (b), or playing (p)							
Merit							
Is structure complete?							

Notes: (What is lacking, etc.) .....

Examiner.....

RECORD SHEET

recorded at the top of column 1. Sixty seconds after giving the child the blocks he is asked what he is going to build and his answer recorded opposite "purpose." If the child sticks to his original intention until his structure is complete or until the time expires he is graded high in the space left for "Dominance of purpose." The following words, or numbers corresponding, may be used in the grading: excellent = 5, good = 4, fair or average = 3, poor = 2, very poor = 1, no purpose = 0. The symmetry of the structure is graded using the same scale, no discernible symmetry equaling zero. Simple structures that give little opportunity for a symmetry, such as wagons, are generally to be graded "average" or "3." In the space left for interest is to be placed "b" or "p," according as the child is interested in building or playing aimlessly with the blocks. The large wheel shown in Plate I has been included in the material for the sake of having something that serves as a distraction from building and suggests playing. It does not lend itself readily to appropriate use with the other material, though, as Plate VIII shows it may be used with very excellent results. The merit of the structure built is determined by comparing it with stereoscopic photographs of structures which have been carefully graded, as will be explained. In the bottom space of the first column is recorded the child's statement as to whether or not his design is complete. As many columns 1, 2, 3, 4, 5, are used as there are attempts at building. It seldom happens that more than five columns are needed for a test. Notes as to causes of incompleteness of structures, attitudes of the child toward the work, his interests, his fixed ideas inhibiting variation and richness of conception, his dexterity, and his type of imagination are recorded in the foot-note or on the back of the record sheet.

The most important, because the most defective, element in the record is the grading of the product for merit as a structure indicative of constructive ability. In order to grade such varied structures a scale of forty stereoscopic photographs has been worked out. Ten judges have contributed to the evaluation of the merit of the elements of the scale. Two students and the writer ranked 250 photographs of structures built by normal and subnormal boys. From this number 54 were selected on the basis of their location in the entire series and given to eight judges, members of the education and engineering faculties of the Uni-

versity of Texas, to rank. The percentage of judges ranking one structure (stereoscopic photograph) above another determined the distance apart in the scale of the two structures. The general method followed in making these calculations is not here described, but is the same as that used by Thorndike in deriving his handwriting scale.¹

From this scale of 54 samples certain ones were dropped because they were types of structures not needed in the part of the scale corresponding to their values. The resulting scale consists of 19 samples of buildings of various sorts, mostly houses, 12 of vehicles, and 9 of machinery, including aeroplanes and the like. The reliability of the scale is high. The grades given to the samples would correlate with other scale values, derived in a similar manner from the judgments of other equally competent judges to the extent of .98.

The probable error of judgment when a single judge grades a single structure is approximately 6 of the scale units, which is about $\frac{1}{8}$ of the range covered by the scale. Expressed in another way, the correlation between an individual's placement of a large number of samples and the true placement in the scale is .92. Practice in grading as few as 40 or 50 samples materially increases the accuracy of grading, decreasing the probable error of judgment by as much as one-half.

¹EDWARD L. THORNDIKE. *Handwriting*. Teachers College Record, Vol. XI, No. 2, March, 1910.

A minor variation from Thorndike's method is in the weighting and averaging of various measures of difference. If $a, b, c, d, e, f, g, h, i$ represent samples in decreasing order of merit the distance between samples e and f may be calculated in nine different ways considering the neighboring four better samples a, b, c, d and the four poorer g, h, i, j . The percentage of judges ranking e above f gives one such measure of difference D_{e-f} . The percentage ranking d above f gives the difference D_{d-f} and the percentage ranking d above e gives the difference D_{d-e} . $D_{d-f} - D_{d-e} = D_{e-f}$. In a similar manner seven other values of D_{e-f} may be obtained. The probable errors of these differences were calculated taking, as a sample the rankings of twenty structures located near the middle of the scale and found to be proportionate to the numbers given below:

		Weights
P. E. of D_{e-f}		
" " $D_{d-f} - D_{d-e}$ or of $D_{e-g} - D_{f-g}$		.1916 5
" " $D_{c-f} - D_{c-e}$ " " $D_{e-h} - D_{f-h}$		.3152 3
" " $D_{b-f} - D_{b-e}$ " " $D_{e-i} - D_{f-i}$		.4092 $2\frac{1}{2}$
" " $D_{a-f} - D_{a-e}$ " " $D_{e-j} - D_{f-j}$		.4740 2
		.5034 1.9

These nine ways of calculating the single difference (D_{e-f}) were averaged, after first weighting as shown, inversely as their probable errors, to give the scale measure separating the samples e and f .

The lowest structure in the scale was originally intended to be a structure of zero value. It was obtained by tumbling the blocks out of a box and selecting the combination that looked most like a structure of some sort. In reality it required no ability to build, but since it looks as though it might be something it seems to have a little accidental merit. It accordingly is not graded zero, but 2, which is the writer's judgment of the distance it lies above a structure of genuine zero merit. This sample is the only one actually built by a boy. Plate III shows it and the one next higher in the scale. Having the scale it then became necessary to establish norms that results might be interpreted intelligently. In doing this the question immediately arose as to whether past experience of various sorts would lead to greater proficiency with the test material.

NO NAME, MERIT 2. HOUSE, MERIT 14
PLATE III.

Thirty-seven children, selected because of experience with metal building outfits, Tinker-Toy blocks, or in manual training, showed an average superiority of nearly a year, at age 14, above children not so experienced. Substantially all of this difference was contributed by children having had, or taking, manual training. There was no evidence that experience with metal building outfits and Tinker Toy blocks helped in the Constructive Ability test. As it is reasonable to think that those who take manual training in the high school are a group selected in large part because of their interest and former ability along

mechanical lines there is no certain, or even probable, evidence that success in the test is not determined solely by native interest and ability. Considering alone the group who had had or were taking manual training it is evident from the data at hand that they stand nearly two scale units above other high school boys of the same age. From *a priori* grounds, based upon similarity of function there seems less reason to think that elements in manual training work would transfer to the test situation than would elements from play with metal and Tinker Toy outfits. As there is no evidence of transfer in the latter case it is added evidence that the difference mentioned, of two scale units, is due to the selection of a group natively more capable along mechanical lines and not due to training. Because of these considerations the norms so far as at present drawn up are based upon the records of all boys examined and not of those only who had had no preceding experience in manual training. Approximately one-quarter of the fourteen year old group were taking manual training when tested. It is not popularly thought that high-school boys are superior to non-high-school boys in mechanical ability. If, however, their constructive ability should differ from that of boys in general the norms here given would on that account be those of a select group.

Three groups of school children, proclaimed "normal" by the principal, were tested, giving the following results:

36 boys, avg age 8.08, built structures of an avg. merit of 44.9,	
	P. E. of avg. = 1 23
33 " " " 12.05, " " " " " " "	58 0
	P. E. of avg. = 1.39
33 " " " 14 23, " " " " " " "	60 8
	P. E. of avg. = 1 47

Boys of ages eight and twelve were chosen to give points for the upper limits of imbecility and moronity, while the age 14 was chosen to carry the scale into the higher ages. The three points shown by stars on Plate IV are the points thus experimentally determined. The curve shown was drawn through these three points to give norms for other ages. It, of course, is an approximation, to be supplanted by actual test records as soon as they are available. The three points determined have considerable reliability as their small probable errors indicate. These three

PLATE IV.

points are sufficient strongly to indicate a curve of the type generally found in the development of intellectual capacities. It is a typical growth curve with the greater part of the increase coming before adolescence. The points at which the curve cuts the ordinates for the different ages give the tentative norms for these ages. These values are given in the following table:

Age	4	5	6	7	8	9	10	11
Mean merit in constructive Ability test	22.4	27.6	32.9	38.8	44.4	49.2	52.9	55.8

Age	12	13	14	15	16	17	18	Adult
Mean merit in constructive Ability test	58.0	59.5	60.5	61.3	62.0	62.6	63.1	63.4

The probable deviation of a normal boy's record from the average for his age is about 7.5 units; 7 for 8-year olds and 8 for 14-year olds. It may be expected that the mean grades given for ages four to six and seventeen to adult, are the least reliable.

PLATE V. MERIT 46.

PLATE VI. TABLE, MERIT 57.

PLATE VII. WAGON, MERIT 61.

Plates V, VI, and VII show the samples in the scale which are closest to the average merits of normal eight, twelve and fourteen-year-old boys. Plate VIII is the scale sample of greatest merit. It is of such worth that structures as good as it may be expected about once out of every 60 built by normal fourteen-year-old boys.

As an illustration of the use of the test and scale, mention may be made of an examination of 250 of the boys in the Texas State Juvenile Training School for delinquent boys. In addition to the idiosyncrasies revealed, the average capacity of the boys was found to be very much less than that of normal boys of the same ages. These boys, of average age 15.26, showed an average accomplishment equal to that of the normal 9.5 year old. The average merit of the structures built by these boys is 51.3, approximately represented by Plate IX. As evidence that the test is highly correlated with mentality it may be reported that these boys showed nearly as great backwardness in such other mental tests as Binet and completion tests.

In addition to difference in merit of structures built the test reveals material qualitative differences. Questioning the boy when building as to what different parts of the structures are for, enables the experimenter to discover much as to the mental processes involved. Young children are prone to build skeleton or incomplete structures and supplement their shortcomings by an active imagination. A house such as that shown in Plate X is typical. With a young child the imagination is rich enough to be adequate for all needs and there is no feeling of incompleteness as to the structure. An older boy would be likely to observe that you would get rained on if you lived in that house. Another characteristic of the products of young children is that their buildings are usually representations of things as they are statically—their ideas frequently do not involve the idea of motion; for example, such a child built the wagon shown in Plate XI, with wheels pegged to the body and incapable of movement. The structure is a wagon without a doubt; for does it not have wheels and a body of a sort? The refusal of the wheels to turn need only worry an unfortunate adult who has not sufficient mental versatility to adjust such minor details.

PLATE VIII. STEAM ENGINE, MERIT 93.

PLATE IX.

PLATE X. HOUSE.

PLATE XI. WAGON.

PLATE XII. BROKEN AUTO IN A GARAGE.

PLATE XIII. AUTOMOBILE

Plate XII, showing a "Broken auto in a garage" was built by a small boy who very prettily gave evidence of several typically childish tendencies. The garage and the automobile are mere skeleton structures, withal quite sufficient to serve as nuclei for fancy to add to. The idea that an automobile is meant to move did not occur to the boy, for it was not until pointed out to him that he saw that it was too big to go into the door of the garage, but at this point a truly liberal imagination came to the rescue for, in reply to the question, "how does the automobile get into the garage," he nonchalantly lifted the garage up and said "Oh, you just put it down that way," replacing it over the automobile.

Plate XIII is an automobile, as any child can prove, since it has wheels, a hood and a steering wheel. True the location of the steering wheel upon the rear axle might be considered slightly unique if one were to stop and think about it, but such reflection is quite unnecessary.

Plate XIV is an illustration of very good manipulative ability coupled with inadequate conceptual processes. The boy started to build a house. Becoming interested in the long rods, beams and wheels he continued to ramble on, building solidly and symmetrically. Lack of further similar material stopped him. When asked to name the structure there was some hesitation—his original intention apparently having been nearly forgotten. After a brief conflict as to whether it should be a house or a bridge he cut the Gordian knot and satisfied all demands by calling it a "house-bridge." It seems evident that this boy worked entirely devoid of ultimate purpose, though quite rigidly imposing upon himself the immediate purpose of making the structure symmetrical. Plenty of other material was available, but he "had to" stop for lack of wheels, rods and beams.

At an older age structures are very likely to be things that work—the wheels of wagons are on axles. Great care may be taken in building an aeroplane that the seat of the aeronaut is so enclosed "that the wind will not sweep him off as his machine rushes through the air."

Plate XV shows a racing automobile, built by an adolescent boy. The wheels turn, the gasoline tank is pegged on, but, keeping in mind that this machine is to travel at great speed it is apparent that the top part of the hood, due to its momentum,

PLATE XIV. HOUSE-BRIDGE.

PLATE XV. AUTOMOBILE.

will fly off from the bottom in taking a sharp turn. The impossibility of pegging the two parts of the hood together to obviate this difficulty was the cause of much worry on the boy's part. The mental processes at work in the building of this racing car were entirely different from those resulting in the preceding five structures. In the main the building of adolescents is of this latter type. Eight year old boys build structures of both types, sometimes of neither, and frequently combine the elements of both in one. The discovery of spontaneous mental processes is one of the most fascinating and, withal, valuable results of the test. The interests of the children, their persistence, the mental restrictions they place upon their own procedure, and their feelings of satisfaction or exasperation with the results of their labors, can hardly be expressed in standardized objective terms, but they are all traits that the test enables the experimenter to observe.

Any number of problems are suggested by the test—the determination of sex differences; the correlation, expressed in terms of coefficients of correlation between the measures of the function here tested and other mental tests; the correlation between success in the test and vocational fitness, etc., etc. The test is particularly applicable in examining children of very different environments, for the trait measured is a creative and executive one, but slightly influenced by the ordinary past training of an individual.

For the present the writer will personally handle the material involved in this test. He will greatly appreciate records from other experimenters using it, that more adequate norms may be rapidly built up and that the test may be used more intelligently in the diagnosis of the innate fitness of youths to engage in mechanical vocations.