1949

The Durost-Fielding Standard Plug Test construction and validation

THE DUROST-FIELDING
STANDARD PLUG
TEST

HAROLD FRANCIS FIELDING
B.S. in EDUCATION

BOSTON UNIVERSITY

SCHOOL OF EDUCATION

Thesis

THE DUROST-FIELDING STANDARD PLUG TEST CONSTRUCTION and VALIDATION

Submitted by

Harold Francis Fielding

(B.S. Boston University 1940)

In partial fulfillment of requirements for

the degree of Master of Education

1949

First Reader:

Walter N. Durost, Associate Professor of Education

Second Reader:

William C. Kvaraceus, Associate Professor of Education

Third Reader:

John G. Read, Associate Professor of Science Education

THE DUROST-FIELDING STANDARD PLUG TEST

CONTENTS

[illegible]

INTRODUCTION

The *Durost-Fielding Standard Plug Test* is a rearrangement type test. The apparatus consists of ten standard plugs resembling those used in industry for the purpose of controlling the sizes of holes reamed to industrial standards. The test calls for the rearrangement of these plugs in the correct order of size. Because of the close limits of precision between the diameters of the different plugs of the apparatus the test may become a means of selecting personnel in industry. Personnel capable of detecting the differences in the diameters of the plugs would most likely be capable of satisfactory work where great precision is required.

Whether the *Durost-Fielding Standard Plug Test* will be of service to industry is a matter to be proved in industry. In a like manner whether such a test can be used advantageously in a school shop situation is a matter which will have to be proved in a school shop. The purpose of this thesis is the development of the original specifications for the test, and the tentative application of the test to school shop problems. However, it is not possible to disassociate this test from the problems of industry because the test is being adapted to select more efficiently those boys who, while working in the school shops, show promise of becoming the more proficient workers of industry.

The *Durost-Fielding Standard Plug Test* was designed, and the specifications adjusted, for use in a special type of shop that includes lathe work, forging, moulding molten metals, sheet iron work, cold iron work, tin and copper-smithing. It is a shop known

[illegible]

[illegible]

[illegible]

[illegible]

as a general shop. The shop is in part-time use by ninth grade classes and equipped for boys on that level of work. The experimental groups are two junior high school opportunity groups who are being trained to adjust themselves to work on a level leading to their initial adjustments to industry.

On its present level of development the Durost-Fielding Standard Plug Test shows promise. With the present experimental groups it is indicated that the test may be used successfully in measuring aptitudes for the work of the shop and in mechanical drawing. A rather high and interesting relationship between the boys' mental ages and ability to score on the test exists, and this measure may be of considerable importance.

Just how the mind and four of the senses are applied to the problem of rearranging the standard plugs in the proper order is a matter of speculation. The sense of sight may be used in rearranging the plugs, but the difference in the diameters is so slight that a very limited success may be achieved through this means. Sight, however, does have a part to play in the more effective use of the gauge which is included with the test.

The sense of hearing also has a limited use, and this sense with the sense of balance seems to operate in aiding the more successful testees. The high scores are made through the sense of touch, and may be made through the sense of touch alone.

How all of these senses are related to mechanical ability and intelligence through scores on the Durost-Fielding Standard Plug Test is the subject matter of this thesis. The test is

definitely non-verbal and suitable for administering to those of low levels of mental ability.

CHAPTER I

STATEMENT OF THE PROBLEM

The present writer has felt for a long time that there was a need for a simple non-language test to get at some of the factors peculiarly related to success in shop work. It seemed that the proposed test should measure more of the native abilities possessed by the testee than the ordinary tests of mechanical ability measure. The test should get somewhat closer to the real nature of the testee and in doing so should involve more of his senses to the uttermost of their relative keenness.

Can it be done.-- True, the mind of the testee, and more, is involved in all tests. What of the infinitely sensitive nervous system related to the sense of touch? No mechanic without a sensitive sense of touch can ever produce the most accurate work. Yet none of the better known tests of mechanical ability measure it. The *Durost-Fielding Standard Plug Test* does measure the ability to feel minute differences in the diameters of its series of standard plugs. What of the sense of sight? The mechanic's sense of sight must be keen and often-times supplemented by the sense of touch in order to assure himself that the results he desired have been achieved. The sense of sight and the sense of touch both supplementing each other are used in detecting the minute differences in the diameters of the plugs of the *Durost-Fielding Standard Plug Test*.

Other senses are involved..- The sense of hearing is involved in much of the work of the mechanic. His sense of hearing is always alert for sounds or changes in sounds which tell him what is going on in regard to his work. His work, then, is based upon his ability to see well, to be able to judge well through the sense of touch, and to gather and act upon information received through the sense of hearing. In the good mechanic all of these senses are keen, alert, and experienced. Where one sense is less keen it may always be taken for granted that the other senses may become supplementary to a greater degree. To some extent all three of these senses play some part in the rearrangement of the plugs of the Durost-Fielding Standard Plug Test.

The sense of balance..- This sense is more or less taken for granted and is somewhat obscure in common thought. The University of Minnesota studies of mechanical ability did not ignore the sense of balance, but finding the correlations low recorded them as such. However, in order to manipulate the gauge properly while detecting the differences in the diameters of the plugs of the Durost-Fielding Standard Plug Test it is necessary to keep the gauge square with the plugs in their vertical positions. This means that the gauge must be kept horizontal during the action of measuring and any deviation from that horizontal position will be the cause of errors in estimating the differences in the diameters of the plugs. The importance of the sense of balance is open to investigation.

What do other tests of mechanical ability measure..- According to J.R.Berkshire and others the test of mechanical ability in the

greatest use in guidance centers today is the Minnesota Paper Form Board.[1] Through the fitting together of many varied geometric forms this test measures the testee's ability to discern from among many shapes those which fit within a specified form. The Minnesota Paper Form Board has been found to have a relationship to the quality of work that the testee is able to produce. This relationship to the quality of the work has been well established as a result of the University of Minnesota studies of mechanical ability.[2] A correlation of .52 between the Minnesota Paper Form Board and the Quality Criterion of that extensive study known as the Minnesota Mechanical Ability Tests[3] has a predictive value of nearly fifteen per cent better than chance alone will give. No one in his right mind will lean too heavily upon a prediction of success based upon this instrument of measure alone. Other tests are needed.

The interest analysis tests.- One of the results of the University of Minnesota studies was in the substantiation of the common knowledge that interest in any work plays an essential part to successful participation. The correlation between their Interest Analysis Blank[4] results and the Quality Criterion of the Minnesota Mechanical Ability Tests was obtained as of .64. This correlation having a predictive value of twenty three per cent better than

1/ J. R. Berkshire and others, "Test Preference in Guidance Centers" Occupations (March 1948) p. 339.

2/ Donald G. Paterson and others, Minnesota Mechanical Ability Tests. The University of Minnesota Press, Minneapolis. 586 pages

3/ Donald G. Paterson and others, op. cit., p. 238.
4/ Donald G. Patterson and others, op. cit., p. 238.

chance alone is considerably better than the Minnesota Paper Form Board predictive value of fifteen per cent better than chance in relation to the same Quality Criterion.

In this investigation considerable thought was given to the interest factor. And accordingly, as one may note further along in the chapter on test procedure, precautions were taken so that general interest would be secured and maintained during the experimental period. Interest in the work, then, being one of the greater factors, one may be justified in ignoring, to a certain degree, the results and findings in regard to many of the tests of mechanical ability where interest appears to be present and the urge to learn and work in the area of mechanics is great.

Tests of mechanical ability are not held in high regard by many personnel workers. The best of the tests of mechanical ability are being reviewed here for what they are worth. In the present writer's opinion, based upon his own experimental data, if one uses several of the best of the tests reviewed in this first chapter, worth-while results may be achieved. Selection, however, would have to be made from those testees who averaged among the top nineteen per cent of those tested. And, as a result of this top selection, none of that nineteen per cent would, according to our experience, be likely to fall below the median of an already carefully selected group of apprentices in training for the machine tool trades.

What is the matter, then, with these tests?..- There is a drawback to the following of such a procedure as that of selecting from the top on such a limited scale as the upper nineteen or twenty per

cent. Rejecting about thirty per cent of the above average applicants, and having them succeed upon being admitted to another school, or after being accepted in another position, is a cause of considerable concern. Of the lower nineteen or twenty per cent who did not rise above the median of success at the school, in one of our experiments, perhaps half of them would have been satisfied with a mediocre success. And according to one educator of national repute, in reference to that particular experiment, be kept from a lifetime of crime.

About half of that lower group of nineteen or twenty per cent were a considerable burden upon their instructors, dissatisfied with themselves in relation to their work, and generally lacking in interest. No testing technique is needed to distinguish from among others this troublesome percentage, but perhaps some other tests could have been used to indicate, in the general direction, just where they did belong. The correlation of .53 between the battery of tests used in the above-mentioned unpublished experiment and the broad and comprehensive criterion of success at the school is the same as that between the Minnesota Paper Form Board and the Quality Criterion of the Minnesota Mechanical Ability Tests.[1/]

What other aspects are there in regard to these tests?..- In order to provide a proper background and understanding it has been necessary to outline in terms of selection and rejection of applicants just what a correlation of .53 entails. In attempting to conduct a

1/ Donald G. Paterson and others, op. cit.

business, in a competitive world, a definite edge over competitors could be acquired if in the initial rejection of prospective employees the elimination of the lower, less efficient, half could be assured. A test, or battery of tests, which would correlate to a reasonably high degree with a broad enough criterion of success related to the trade or business would be necessary. This is where tests of mechanical ability, if used singly, fail. The tests if used singly are not broad enough to contain within them any real measure of an extensive criterion. If an extended battery of the best tests is used, then, some reasonable results may be obtained. According to our experience, the correlations of each test with a broad criterion will be lower as the criterion becomes broader, but some of the testees will be found at the top of most all of the tests, and their high averages may be computed. These testees may be selected with some assurance that better than average testees have been selected.

What are these better tests which may be used in a battery?..- According to J. R. Berkshire and others, and in the order given in Test Preferences in Guidance Centers[1/] they are:

1. The Minnesota Vocational Test for Clerical Workers
2. The Minnesota Paper Form Board
3. The Purdue Peg Board
4. Mechanical Comprehension Test, Bennett and Fry
5. Minnesota Rate of Manipulation
6. O'Connor Tweezer Dexterity

1/ J. R. Berkshire and others, op. cit. p. 339.

7. O'Connor Finger Dexterity

8. MacQuarrie Test for Mechanical Ability

9. O'Rourke Mechanical Aptitudes Tests

10. Minnesota Mechanical Assembly

11. Detroit Mechanical Aptitudes Examination for Boys

12. O'Connor Wiggly Block

These are the tests in greatest use today. What is it that they do not measure that we expect the Durost-Fielding Standard Plug Test to measure?-- The name of the test does not always reveal the exact uses to which the test may be put. The personnel worker discovers this and his skill in the use of tests is the result of extensive study and experience in relation to the particular job that he may be doing. For example, why is the Minnesota Vocational Test for Clerical Workers at the head of the list? It may not belong there but it is a good test for mechanical ability. We will, however, have to quote a noted test expert, Walter van dyke Bingham,[1] to justify this position. "Certain tests have been christened with names implying that they measure abilities other than those they actually sample. A few, curiously, have been found to be truly significant in ways which their inventors did not suspect. MacQuarrie's Test for Mechanical Ability (so-called, whatever may be the combination of perceptual and manipulative processes which it samples), has been found by at least one investigator to correlate with subsequent progress in office work better than do certain tests designed

1/ Walter van dyke Bingham, Aptitudes and Aptitude Testing, Harper & Brothers Publishers, New York, 1937. p. 9.

to measure clerical aptitude. Similarly, the number-checking and word-checking test, known as the Minnesota Vocational Test for Clerical Workers, has been found to correlate better than the MacQuarrie with measures of progress of toolmaker apprentices! What's in a name? Danger and confusion, if the label misleads anyone into using a test without first assuring himself as to two things: First, the nature of the specific abilities exercised by an individual when taking the test; and second, the need of the same abilities when learning to do the work of the occupation in question. The Durost-Fielding Standard Plug Test will easily meet one requirement here; the sense of touch and the sense of sight is universally used in the work of the mechanic.

Resemblance to job performance not a factor..- Resemblance of performance on the Durost-Fielding Standard Plug Test to most of the work of the mechanic is very limited. Many tests of mechanical ability contain resembling features of the mechanic's work, and the limitations in regard to what they measure are many." The Durost-Fielding Standard Plug Test, we hope, measures more of the individual and less of job resemblance, and we think we have evidence of a correlation with subsequent success.

The Purdue Peg Board[1/]..- This test, like the O'Connor Finger Dexterity Test,[2/] measures the rate of manipulation of small objects.

1/ Joseph Tiffin, Industrial Psychology, Second Edition, Prentice Hall Inc., New York, 1947. pp. 126-7, 141.

2/ Johnson O'Connor, O'Connor Finger Dexterity Test, Worksample No. 16. Human Engineering Laboratory, Stevens Institute, Hoboken, New Jersey.

The O'Connor Tweezer Dexterity Test[1] measures the rate of manipulation with small tools. It remains to be seen just how the sensitivity of touch measured by the Durost-Fielding Standard Plug Test correlates with the rate of manipulation of small objects.

The Mechanical Comprehension Test..- "This test contains sixty pictorially presented mechanical problems. A correct solution to each question requires the understanding of the operation of physical principles rather than rote knowledge... This type of test measures a mechanical ability of a higher level than do most tests of mechanical ability. It does not test mechanical assembly performance or manual dexterity."[2] The Mechanical Comprehension Test measures many of the physical principles which would be within the understanding of a mechanic. In industry these principles would be understood as included within the meaning of the general term know how. There is a certain amount of know how in the way one goes about selecting, arranging, and measuring the plugs of the Durost-Fielding Standard Plug Test. It may be that one who knows how to go about the doing of things with an understanding of the physical principles involved would have some of the abilities needed to score successfully in the measuring and arranging of these plugs.

1/ Johnson O'Connor, O'Connor Finger Dexterity Test, Worksample No. 16. Human Engineering Laboratory, Stevens Institute, Hoboken, New Jersey.

2/ George K. Bennett and Ruth M. Cruikshank, A Summary of Manual and Mechanical Ability Tests. (Preliminary Form) The Psychological Corporation, 522 Fifth Avenue, New York, N.Y. 1942. p. 37-38.

The Minnesota Rate of Manipulation Test..[1/] The apparatus for this widely used test consists of a board containing sixty circular holes into which fit sixty smaller cylinders. The blocks may be transferred from a table to the holes in the board or in another test they may be lifted out of their holes with one hand and turned over with the other hand. The Durost-Fielding Standard Plug Test. WHILE ITS APPARATUS CONSISTS OF CYLINDERS AND HOLES IT IS NOT A MANIPULATION TEST AND TIME DOES NOT ENTER INTO ITS SCORING. Rate of manipulation cannot be measured through the Durost-Fielding Standard Plug Test. Mechanics who take naturally to the kind of work that the Durost-Fielding Standard Plug Test represents would resent time-pressure work. When any work or performance takes patience it also takes time. The testee must never know he is being timed on the Durost-Fielding Standard Plug Test and time is not mentioned. Should the testee inquire as to whether or not he is being timed he should be told that he is not. And in the interest of real honesty it would be just as well if no stop watch is held on the test. The average time taken by testee is seven and one-half minutes. When the test was timed the stop watch was concealed.

The MacQuarrie Test of Mechanical Ability[2/]..- The MacQuarrie Test is a battery of seven sub-tests. The first sub-test, called tracing, measures hand and eye co-ordination. The second sub-test,

1/ W. A. Ziegler, Minnesota Manual Dexterity Test; Minnesota Rate of Manipulation Test. Publisher, Educational Test Bureau, Minneapolis, Minnesota, 1931.

2/ T. W. MacQuarrie, The MacQuarrie Test of Mechanical Ability, The California Test Bureau, 1925.

called tapping, putting three dots in a series of circles as fast as possible, measures the motility of the pencil-holding hand. The third sub-test, called dotting, measures precision or aiming, and is limited to measuring the co-ordination of the hand, pencil, and eye. The fourth sub-test, called copying, involves judgment in regard to space, and form. The fifth sub-test, called location, involved a sense of space without regard to form. The sixth sub-test, called blocks, measures thinking in the third dimension. The seventh sub-test measures the ability of the eye to trace lines through a maze-like pattern and perhaps the ability to follow lines through a complex drawing such as a mechanic may have to read. The MacQuarrie Test for Mechanical Ability is a rough measure of the individual. And as Walter van dyke Bingham writes:[1/] "MacQuarrie's Test for Mechanical Ability (so called-- whatever may be the combination of perceptual and manipulative processes which it samples), has been found by at least one investigator to correlate with subsequent progress in office work better than do certain tests designed to measure clerical aptitude." Here, again, we seem to be measuring the individual. As with the Durost-Fielding Standard Plug Test, which does not resemble either mechanical drawing or the work being done in the school shops in the most remote way, the test measures the individual. We think that the Durost-Fielding Standard Plug Test indicates that the individual tested has to relative degrees, based upon the keenness of his senses, not

1/ Walter van dyke Bingham, op. cit. p. 9

mechanical ability, or clerical ability, but just the ability to alert and apply his native endowment in relation to problems involving sight, touch, and hearing.

O'Rourke Mechanical Aptitude Test-Junior Grade..- Part I has twelve sub-tests. For each of six groups of pictorial illustrations the student is to match corresponding tools and objects. Another part requires the examinee to identify the use of the tools and objects which he has just matched. Part II is a series of sixty statements on mechanical information. Answers are supplied in multiple-choice form. The test is limited chiefly to material of the information type. "While there is likely to be some correlation between knowledge of mechanical tools and mechanical aptitude, the correlation is by no means perfect."1/ The O'Rourke Mechanical Aptitude Test,2/ depending almost entirely upon the testee's information and his ability to read, is the sort of thing we are avoiding in the non-verbal Durost-Fielding Standard Plug Test. There are enough tests of the O'Rourke Mechanical Aptitude type. Similar tests like the Detroit Mechanical Aptitudes Examination for Boys3/ contain sub-tests which bear a resemblance to the features of many other tests. The recombining and extending of the forms of the existing tests of mechanical ability has by no means reached the

1/ George K. Bennett and Ruth M. Cruikshank, op. cit. p. 50

2/ L. J. O'Rourke, O'Rourke Mechanical Aptitude Test-Junior Grade. The Psychological Institute 3506 Patterson St., N.W., Washington, D. C.

3/ Harry J. Baker and A. C. Crockett, Detroit Mechanical Aptitudes Examination for Boys. Public School Publishing Co., Bloomington, Illinois.

saturation point. Some tests measure on the higher levels of mechanical ability and these contain measures of knowledge in regard to physical science, mechanical drawing, and arithmetic. The Mechanical Comprehension Test[1] MEASURES ON THIS HIGHER LEVEL, while the MacQuarrie Test of Mechanical Ability measures better in the LOWER PERCEPTUAL AREAS of mechanical ability.

Visual construction tests..- The Johnson O'Connor Wiggly Block Test[2] is a block of wood cut into nine wavy-formed pieces. The testee assembles the nine pieces as quickly as possible. In using such a test one assumes that the test measures the ability to perceive the sizes, shapes and relations of objects and the relation of those objects to each other. If this is done quickly, efficiency in visual construction is credited to the testee. Another test of this nature is the Crawford Tridimensional Structural Visualization Test.[3] The apparatus for this test is an aluminum base board with projections and depressions into which nine irregularly shaped blocks fit to form a circular block. In reconstructing the cylinder the score is the time necessary to complete the assembly. These tests are not tests of general mechanical ability but tests which measure some part of mechanical ability just as over a hundred other different tests measure some other part of that ability known

---

1/ George K. Bennett, Mechanical Comprehension Test. The Psychological Corporation, 522 Fifth Avenue, New York, N.Y.

2/ Johnson O'Connor, Wiggly Block, C. H. Stoelting Co., 424 N. Homan Ave., Chicago, Illinois.

3/ J. E. Crawford and D. M. Crawford, Crawford Tridimensional Structural Visualization Test, The Psychological Corporation.

as mechanical.

Assembly tests..- The ability to assemble such devices as locks, light sockets, pliers, and other more or less common articles of commerce known as hardware is thought to be a reliable measure of mechanical ability. The better assembling tests for general mechanical ability are the Minnesota Assembly Test and the Stenquist Assembling Test of General Mechanical Ability. The Minnesota Assembly Test 1/ is a modification and amplification of the Stenquist Assembling Test of General Mechanical Ability. 2/ While these tests are our best, they are cumbersome and require considerable attention. Essential parts of the tests tend to disappear as a result of the actions of some of the testees and are sometimes very difficult to replace. These tests measure over a wide area of mechanical ability including structural visualization, mechanical comprehension, experience, interest, and elements of sight and touch. Our highest correlations with general mechanical ability have been obtained through the use of these two tests. Correlations of .90 between this type of test and shop work have been considered by many workers in this field as being the ultimate of what may be achieved in measuring mechanical ability.

Is the field open for yet another test?..- Exploration in this field has been considered by many to have reached a point of

---

1/ Donald G. Patterson and others, Minnesota Mechanical Ability Tests. The University of Minnesota Press, Minneapolis, Minnesota. 1930 pp. 102-3.

2/ J. L. Stenquist, Assembling Test of General Mechanical Ability. C. H. Stoelting Co., 424 Homan Ave., Chicago, Ill.

saturation. However, some of the more advanced workers find that the field of mechanical ability is still an inviting area for investigation, and in conflict with general opinion some areas of mechanical ability are relatively unexplored. Any advance in the field of testing for mechanical ability, however, will come only as the result of many more investigations and experiments, and a greater command of, and better organization of the knowledge and experience gained by others. In such a direction, and after this manner there will be those who will eventually construct the better tests.

Precursors of the Durost-Fielding Standard Plug Test.-- While it is considered that there is no validated test extant which measures in the area of mechanical ability now being measured by the Durost-Fielding Standard Plug Test, precursors of the test may be recognized readily in the work of the skilled mechanic. Saving time and eyesight the mechanic may set his calipers to the size of a standard plug using a feeler of some known thickness where oversize turning is required. This act and the act of measuring his work in the lathe resembles the performance of measuring with the gauge of the Durost-Fielding Standard Plug Test. The gauge used with the Durost-Fielding Standard Plug Test is quite similar to those gauges known as snap gauges which are used for many purposes of measuring and testing throughout the machine tool industry.

A precursor on the psychometric test level..- As reported by Dr. Walter N. Durost of the faculty of the School of Education, Boston University, Albert G. Packard, Director of Aptitude Testing

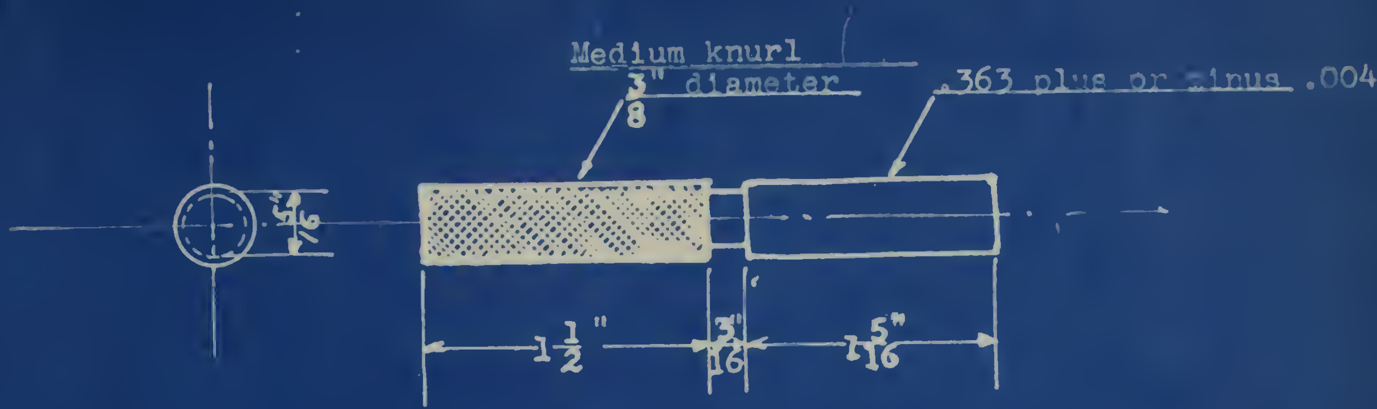
in the schools of Baltimore, Maryland, used a test involving the same principles as those used in the Durost-Fielding Standard Plug Test. No data are available in regard to Albert Packard's work. No knowledge is available anywhere as to the limits of human sensitivity in measuring the differences in the diameters of the cylinders in such a test as the Durost-Fielding Standard Plug Test. It is in this investigation that we proceed to make that information available.

Interest in this test..- It was the persistent interest of Dr. Walter N. Durost in the measurement of the sense of touch in relation to mechanical ability that initiated the proposal to construct and validate the Durost-Fielding Standard Plug Test. While there are many other ways of measuring the sense of touch, this particular type of apparatus was constructed because its plugs and gauge resemble so closely the plugs and gauges already in use in industry. The present writer also believes that rapport with the testee can be more readily established through the use of apparatus which has a familiar look. The fact that more than the sense of touch was being measured is interesting, but this was not apparent until a considerable number of testees had been observed. It was also supposed that the test would get at factors that are beyond the reach of the current tests of mechanical ability and this too was a matter of interest. Then, again, it was thought that in this neglected area of mechanical ability some hidden factor of perhaps considerable importance might be revealed. None of the greater experiments in the field of mechanical ability ever having recog-

nized the existence of this important area—which is so closely related to the work of the machine tool and precision type of mechanic—it seemed possible that virgin territory might be explored.

THE DUROST-FIELDING STANDARD PLUG TEST APPARATUS

DUROST-FIELDING STANDARD PLUG TEST FOR MECHANICAL ABILITY



INSTRUCTION SHEET

Select stock 3/8" drill rod

Cut stock to 3" lengths

Square ends

Center in drawing in chuck

Neck according to drawing

Knurl 1 1/2" length

Turn to .363 plus or minus .004

Heat to 1650 deg.

Quench in lard oil

Ring lap to size and finish

Where a cylindrical grinder is available, harden without turning and grind to specifications, allowing for a distribution of diameters plus and minus .363 or attempt to reproduce the sample test battery plugs.

TENTATIVE TEST BATTERY

SAMPLE OF TEN PLUG DIAMETERS SELECTED FROM A DISTRIBUTION OF PLUG DIAMETERS TURNED BY THE PUPILS OF EASTERN JUNIOR HIGH SCHOOL, LYNN, MASS.

1	.3750
2	.3669
3	.3652
4	.3650
5	.3635
6	.3632
7	.3629
8	.3618
9	.3580
10	.3388

PLUGS TURNED BOTH LARGER AND SMALLER THAN SPECIFICATIONS ARE USED AT THE EXTREME ENDS OF THE BATTERY.

MOST OF THE PLUG DIAMETERS WILL CLOSE TO .363

CHAPTER II

Description and Function

of

THE DUROST-FIELDING STANDARD PLUG TEST

The Durost-Fielding Standard Plug Test.-- This test is a non-verbal test designed for the purpose of measuring abilities in the lower areas of mechanical aptitude. The test consists of ten hardened carbon steel plugs, having a knurled handle and a necked-in middle section. The middle section, while necked in for purposes of manufacture, serves well in containing an identification tab. The handle is knurled to help prevent mistakes in identification of the proper end of the test plug. The handle also serves as a means of holding the plug should a testee wish to remove the plug from the rack for the purposes of measuring it.

In administering the test..- A small rubber band may be used in the necked-in section for covering the identification tab which is found at the bottom of that groove. There is, of course, no intelligible sequence to the plug identification tabs, but a rubber guard must be used in any serious work to prevent a carry-over through traits or acts of memory.

The test gauge.-- A tool known to industry as a snap gauge was designed for use with the test battery of plugs. After some experimentation a distance of .3725 parts of an inch was found to be

suitable for the gap of the gauge. The gap in the gauge has a fixed relationship to the diameters of the standard plugs, all of which are smaller than the gauge except one. This one plug is larger by .0025 of an inch and the gauge does not fit over it. Some testees, as we have found out, do not realize that this plug which does not fit inside of the gauge must be larger. So after much consideration this largest plug was left in the battery.

In experimenting.-- In trying out over forty different plug sizes, the plug sizes submitted in the following order have offered the most promising distribution of test scores.

Largest plug----------A	.3750"	
		----plus.0025"
Gauge gap-------------	.3725"	----
		----minus.0081"
Plug------------------B	.3669"	----
		----minus.0017"
Plug------------------C	.3652"	----
		----minus.0002"
Plug------------------D	.3650"	----
		----minus.0015"
Plug------------------E	.3635"	----
		----minus.0003"
Plug------------------F	.3632"	----
		----minus.0003"
Plug------------------G	.3629"	----
		----minus.0011"
Plug------------------H	.3618"	----
		----minus.0038"
Plug------------------I	.3580"	----
		----minus.0192"
Plug------------------J	.3388"	

There is a critical difference between some of the plugs of this battery which will really test the ability of anyone who attempts to rearrange them in the correct order.

The present specifications..- The set of the gauge, with its present amount of difference between the gauge and the diameters of

the plugs is important. Any change in the dimensions, either in the diameters of the plugs or in the set of the gauge could be made only on the basis of a much greater extension of experience with this type of test. We have no precedent to guide us in this original contribution to knowledge, and will have to wait upon time and perhaps the experiences of others.

The reference strip.-- The aluminum reference strip with the identification symbols stamped upon it serves as a matter of convenience. It must, as a matter of course, always be concealed from the testee. It does serve to save time between tests and in arranging the planned disarray. Its use is optional.

The identification symbols..- The letter or number which identifies each plug may some time need to be changed, but this does not offer any great difficulty. The ten letters or numbers identifying the plugs are typed on a strip of white gummed paper. The paper is then cut in strips 3/16 of an inch wide and two inches long. These strips are then wrapped around the circumferences of the necked-in sections of the plugs. Strips of scotch tape are then cut to a width of 3/16 of an inch and two inches long, and are drawn tightly over the white paper strips, serving to hold them in place and keep them clean.

Administering the Durost-Fielding Standard Plug Test..- First the plugs must be in planned disarray. This is simply that plugs M-H-Y-S-F-D-B-E-C-X are presented to the testee in the order E-D-C-M-X-H-Y-S-F-B. The testee is seated and the administrator stands at the testee's left, facing in the same direction as the

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

testee. The testee is told to rearrange the plugs. The administrator's words are: "I want you to find the largest plug and place it here to the left." The place to the left is indicated by the administrator using the index finger of his own left hand and placing that index finger on the top of the plug in the planned disarray which is at the extreme left-hand end.

The administrator, having now started the process, must continue, saying: "And the next largest plug must be placed along side of it." For each plug, and so on down to the smallest plug, the administrator places the index finger of his left hand on the top of the plug next in the disarray and says: "And this plug will be the next largest one." On reaching the smallest plug the administrator says: "And this plug will be the smallest one." Then indicating the plug to the left, in the direction toward the larger plugs, the administrator says: "And this plug will be larger." And so on until the largest plug is again reached. If the testee shows a little impatience the administrator may speed up the given directions and slide over the tops of several plugs saying only "Larger, larger, larger." If, however, the testee does not nod his head as the finger touches each plug and as the administrator says "Larger, larger, larger," then the full directions should be carried out. The head may not nod, but there may be a slight rhythmic movement of the body which indicates that further directions are anticipated to be the same, and sign of impatience may be indicated by irregular, involuntary movements or sounds. Avoid

irritating the testee. Be complete and go slow with the dull testee. Use good judgment in each case. Watch the performance of the testee, there is much to be learned by doing so.

The administrator then, while taking the gauge in his own right hand, says: "Use this gauge like this," and while making one pass of the gap end of the gauge over the plug which sets in the place where the largest plug is to be, says: "The largest plug will be at this end." The testee is then handed the gauge. There is no time limit, but a record of the time should be kept.

<u>The placing of the larger plugs to the left is important</u>..- The test was designed so that all of the plugs would be moved a maximum distance. Placing the larger plugs to the right would present an advantage, leaving two of the more critical plugs in the planned disarray too close to the correct position in the correct rearrangement.

<u>The passing of the gauge</u>..- All that this passing of the gauge over the plug involves is that for a moment the plug is seen by the testee within the gap of the gauge. The passing movement of the gauge is a straight movement without any oscillation or vibration of the gauge. This passing movement is the simplest, easiest, and shortest movement possible.

<u>No techniques are to be taught</u>..- It is probable that the administrator should have a steady hand. Any unsteadiness, oscillation, or vibration detected by the testee through either sight or hearing could suggest the better techniques to the testee. In the single limited pass made by the administrator there should be no twisting

movement and squaring up or leveling off of the gauge while in contact with the plug. Such action would be teaching technique which the testee should either know himself or discover himself. The administrator also shall not place his finger on the plug to steady it while passing the gauge over the plug. This also is included among the better techniques which shall not be taught. No testee who fails to steady the plugs during the process of re-arranging them can possibly obtain a high score. Therefore, no suggestion of this particular technique or any other is permissible.

SCORING THE TEST

Scoring a rearrangement test..- The rearrangement test is unique among the types of objective tests in that it is concerned with the relations which exist among a series of items rather than in the items themselves. While this is so of all paper and pencil type rearrangement tests it is also true of the Durost-Fielding Standard Plug Test. In a like manner the plugs of this test are not items of importance, but the relationship that exists in a rearrangement of the series of plugs is all that can be of importance. The correct rearrangement of the plugs of the Durost-Fielding Standard Plug Test can be achieved, but that correct order is so difficult to attain that very few persons will achieve a perfect score. The order in which the plugs of the test may be arranged by the testee presents a problem in scoring common only to this type of test.

The value of the difference-scores in terms of chance..- In a rearrangement test such as the Durost-Fielding Standard Plug Test, where chance is a variant, it is possible to arrive at the chance values of the difference-scores. Cureton and Dunlap[1/] proposed the calculation of the rank order of correlation for each testee's individual score. This proposal does not involve too much difficulty in a test of the length and nature of the Durost-Fielding Standard Plug Test.

1/ E. E. Cureton and J. W. Dunlap, "Scoring the Rearrangement or Continuity Test", The School Review (1930) 38: pp 613-616.

However, this proposed method has, according to D. A. Worcester,[1] a great weakness. Large errors, he writes, "...are due to outright guessing and that, as such, does not deserve so heavy a weighting as the process of squaring imposes."

Should a weighted scoring method be used?..- Whether guessing in regard to the Durost-Fielding Standard Plug Test should be heavily weighted by squaring the differences between the correct order of the plugs and the order found in the rearrangement of the plugs by the testee is a matter of interest. In real life guessing in regard to minute differences required in fine measurements and adjustments could be of serious consequences and heavily weighted with failure. To succeed in what perhaps could be important work requires not guessing but exact skills and patience. Cureton and Dunlap[2] may have more to offer through their heavily weighting of the lack of skill, and of guessing, and of the lack of patience than is generally understood. While in this experiment no use was made of the weighting of scores the method should be made available and the process should be understood. In the direction and nature of further research the weighted scores technique should be explored. Table No. 1 (p. 29) enables one to compare the weighted scoring method with variants of several other scoring methods. Judging from an inspection of the available data and from an inspection of Table No. 1 the weighting of our scores

1/ D. A. Worcester, "Still Further Comments on the Scoring of Continuity Tests", The School Review (1930) 38: pp 462-466.

2/ E. E. Cureton and J. W. Dunlap, op. cit.

seems to make very little difference in the comparable results.

How is the weighted score obtained?..- The method is simple enough and some one may wish to use it. The ten plugs of the Durost-Fielding Standard Plug Test in their correct order presents, in this experiment, a rare instance of a perfect criterion. The rearrangement of the ten plugs by the testee may be considered as ten cases or observations and allow us to enter formula (234) of Dunlap and Kurtz Hand Book[1/].

Data needed to enter the formula:

- d represents the difference in rank between the correct order of the plugs and the testee's rearrangement of the plugs.
- N represents the total number of plugs or possibilities to observe a difference in rearrangement.
- $\sum$ represents the sum of and is found in relation to d^2.
- d^2 represents the differences in the placement of the plugs by the testee from the correct order when that difference is multiplied by itself.

The rank correlation coefficient is represented by the Greek letter rho (ρ).

Formula 234, $= 1 - \frac{6 \sum d^2}{N(N^2-1)}$, taken from Dunlap and Kurtz Handbook [2/] is the formula used by Cureton and Dunlap[3/] in

1/ Jack W. Dunlap and Albert K. Kurtz, Handbook of Statistical Nomographs, Tables, and Formulas. World Book Company. Yonkers, New York, 1932, p. 36.

2/ Jack W. Dunlap and Albert K. Kurtz, op. cit. p. 36.

3/ E. E. Cureton and J. W. Dunlap, loc. cit.

their weighting of the scores. The same formula is also being used throughout this experiment because of the limited number of cases involved and because any greater refinement of the calculations is unnecessary.

Dunlap and Kurtz Nomograph 37 [1] may be used instead of Formula 234 and will facilitate the calculation considerably. Nomograph 37 has been used freely to check the calculations based upon Formula 234, and Nomograph 43[2] has been used to estimate the probable error of the corrected rank correlation coefficients. The use of Nomograph 43 for the purposes of estimating the probable error of the corrected rank correlations has been checked by carrying through the calculations on the basis of Formula 238 A. of Dunlap and Kurtz Handbook.

Other references for calculations..- Cureton and Dunlap[3] have provided in The School Review a nomograph similar to Nomograph 37.

Table No. 1 of this experiment..- In Table No. 1 (page 29) it may be noted that in the column of difference-scores the difference-scores are evenly spaced. The column of Sims-scores, derived through calculations from comparable difference-scores, and not by a process where the scores are weighted by the squaring of errors, has also an even distribution of scores throughout the sixteen steps in that column.

1/ Jack W. Dunlap and Albert K. Kurtz, loc. cit.

2/ Ibid, p. 43.

3/ E. E. Cureton and J. W. Dunlap, loc. cit.

Where the rank correlation coefficient is calculated as in the column designated by ρ (rho), and also in the column designated by the per cent of efficiency greater than chance alone (per cent eff.), the sixteen steps are as a result uneven. And on the level of difference-scores 24 and 26, there is an overlapping of scores. This overlapping, of course, would not happen if the median of such scores was selected from a large number of scores calculated by the same process. The mean chance scores, thus obtained, would approximate in evenness the steps of the difference-score and Sims-score columns.

Chance operating through a random selection of scores reveals the degree of coarseness in the method of Cureton and Dunlap,[1] but due to the general reasonableness with which chance operated as shown by the distribution of scores on the sixteen steps none of the scoring methods need be invalidated for the purposes of this experiment.

1/ E. E. Cureton and J. W. Dunlap, loc. cit.

Table No. 1. The relationship between deviation from the correct placement of the Durost-Fielding Plugs and the correct order expressed by several variant methods together with the distribution of scores for 51 individuals tested.

1	2	3	4	5
Diff.-Scores	Frequency of scores	ρ 1/	Sims 2/ Scores	Per cent of eff. better than chance
0		1.000	10.0	100.0
2	sss	.988	9.4	84.6
4	ss	.975	8.8	78.8
5				
6	sssss	.928	8.2	64.3
7	ss			
8	ssss	.915	7.6	61.3
9				
10	sssssss	.805	7.0	42.5
11	ss			
12	sss	.783	6.4	39.6
13				
14	sssss	.748	5.8	35.4
15				
16	sss	.680	5.2	29.3
17	ss			
18	ss	.660	4.6	26.4
19				
20	ss	.640	4.0	24.7
21	s			
22	ss	.616	3.4	22.7
23				
24	s	.362	2.8	7.4
25				
26	s	.381	2.2	8.1
27				
28	s	.275	1.6	4.2
29				
30	s	.220	1.0	2.7
31	s			
32	s	.190	.4	2.0

1/ Jack W. Dunlap and Albert K. Kurtz, Handbook of Statistical Nomographs, Tables, and Formulas. World Book Company, Yonkers-On-Hudson, New York, 1932 pp. 36-37.

2/ Verner M. Sims, "The Scoring of the Rearrangement Test", Journal of Educational Psychology, 28: 1937 pp. 302-304.

Adopting the Sims-score..- Before we adopt the Sims[1] method of scoring a good example of just what we are up against should be given: For example, here is the basis of either a Sims-score or a difference-score. The sum of the deviations squared is the basis for prediction in both cases where the difference-score is 16.

D	D^2	D	D^2
0	0	0	0
8	64	2	4
0	0	2	4
0	0	2	4
0	0	2	4
0	0	2	4
0	0	2	4
0	0	2	4
0	0	2	4
8	64	0	0
16	128	16	32

$$\rho = 1 - \frac{6 \sum d^2}{N (N^2 - 1)} = \rho .22 \quad \text{when } d^2 = 128$$

$$\rho = 1 - \frac{6 \sum d^2}{N (N^2 - 1)} = \rho .81 \quad \text{when } d^2 = 32$$

The predictive value of ρ .22 is 3 per cent better than chance alone, and is not considered significant.[2]

The predictive value of ρ .81 is 43 per cent better than chance alone, and is considered very significant.[3]

See Formula 234 and nomograph 37.[4]

1/ Verner M. Sims, op. cit.

2/ Walter van dyke Bingham, Aptitudes and Aptitude Testing, Harper Bros. 1937 pp. 258-259.

3/ Ibid.

4/ Jack W. Dunlap and Albert K. Kurtz, op. cit. pp. 36-37.

In the case where $d^2 = 32$, such a predictive value of 43. per cent better than chance is quite valid. The differences, or errors in the rearrangement of the plugs are evenly distributed, and a certain degree of skill in which one may place some confidence is assured.

In the case of the extreme differences, where $d^2 = 128$, a retest is certainly to be recommended. This particular test score indicates that a very high degree of skill is being used by the testee, and that the difference of 8 places in the interchange of the two plugs is an error of chance and not evidence of lack of skill. A certain amount of carelessness, or negligence, may be in evidence. In some situations a testee showing evidence of such carelessness, or negligence, could not be looked upon as dependable. A thorough re-examination of such a testee would be in order. And perhaps the predictive value of this extreme but hypothetical case would, after all, be the correct one.

Shall we adopt the Sims scoring method..- Now that we know the worst that can happen, what we want to know is, does it happen? With the possibilities of chance operating in this worst possible way, and our difference-score table constructed in such a way that chance was free to operate, we did not obtain a compilation of predictive values that would lead to any grave error. In the light of our experiences here we do not have to be too much concerned, from a practical viewpoint, with the operations of chance upon our difference-scores.

The method of figuring the Sims scores..- Where (S) = the Sims score, and (N) = the number of plugs in the set, and (d) = the number of places the total of all of the plugs in the testee's rearrangement are out of place, the following formula is given.[1]

$$S = N - \frac{3d}{N}$$

The basis of the Sims score, as seen here, is always the difference score, or the sum of the deviations. In this way the errors are not squared, and as in the cases of our hypothetical scores of (16) the chance errors, or deviations, are not multiplied. Thus high skill with a major error would be in score form represented by an average score of (16), and average skill with no major error would be represented also by a score of (16). Remembering our hypothetical cases ($d^2 = 32$) and ($d^2 = 128$) both would become average scores of (16) and given a Sims score of 5.2, which would compare with any other Sims score of 5.2.

Procedure for setting up the Sims score..- Start with the highest score; the one having the least deviation from the correct order. Figure only those scores where (d) is an even number. Figure all of the scores in order taking the higher scores first. When zero is approached cease figuring. Below zero in the Sims scale of scores, and at (d 32) for difference-scores, no scores are evidence of either aptitudes or skills and are to be considered merely as the products of chance alone.

1/ Verner M. Sims, op. cit.pp. 302-304

According to Conrad,[1] and the tables he has so painstakingly compiled, a score as low as (50) is possible with the Durost-Fielding Standard Plug Test. A score of (33) in Conrad's tables also checks with the Sims scores mean chance deviation, and according to both Conrad and Sims[2] all scores below (33) may be disregarded as the products of chance alone.

1/ Herbert S. Conrad, "The Scoring of the Rearrangement Test". Journal of Educational Psychology 27: Apr. '36, pp. 241-252.

2/ Verner M. Sims, op. cit. p. 303.

CHAPTER III

The Description and Selection

of

THE PROBLEM

Orientation..- A serious attempt is being made in this experiment to find a simple non-verbal test which may be used effectively in measuring mechanical ability or the intelligence related to mechanical ability. While the Durost-Fielding Standard Plug Test may be adapted to the measuring of mechanical abilities and related intelligence among a normal population, in this case the experiment is limited to the discovering of such abilities among mentally retarded boys.

Locale..- The school site of the experiment is in the general shop and the mechanical drawing rooms of Eastern Junior High School, Lynn, Mass. The boys of the experimental groups are boys assigned to the opportunity classes and are at least three years retarded in grade placement. The training in both the shop and the mechanical drawing room is related to any industry where things are measured, squared up, placed in line, assembled, or arranged in order. The direction of the work is toward semi-skilled labor, where the need of acquired knowledge is not great, but where intelligence enough to understand, and practice enough in doing what one is told is an asset.

Critical data on the experimental group..- Some of the boys of the experimental group, while statistically regarded as mentally retarded, show evidence at times of greater intelligence than their classification indicates. One-fourth of the boys are classed as dull normal and borderline cases. One-half of the boys could be classed as morons, and the lower one-fourth of the group as borderline feeble-minded. There are no feeble-minded boys, no imbeciles, and no idiots in the group... All of the boys are capable of satisfactory adjustments to society. The data on the group is as follows:

OUR EXPERIMENTAL GROUP

Critical Data on the Group of Twenty-Eight Boys

	Chronological Age	Binet Mental Age	Binet I.Q.
Upper limit of the range	15 yrs. 6 mo.	12 yrs. 5 mo.	95
3d Quartile	14 yrs. 6 mo.	9 yrs. 10 mo.	79
Median	14 yrs. 2 mo.	9 yrs. 10 mo.	74
1st Quartile	13 yrs. 6 mo.	8 yrs. 8 mo.	66
Lower limit of the range	12 yrs. 0 mo.	6 yrs. 10 mo.	55

The critical data in regard to the boys of the experimental group was compiled from psychological reports which record the progress of the boys. The data is taken as of October 1, 1948.

The problem of developing a test for the opportunity class type of boy..- There are plenty of boys and men around who are not too bright, and among such types, and among others who were known to be bright and even brilliant, the first experimenting was done. The experimental group of twenty-eight boys whose critical data are compiled herein was not contaminated by tentative experimental techniques.

The tentative test..- A test consisting of four plugs having diameters close to three-eighths of an inch, but varying in differences of .0005 parts of an inch, was tried by a number of individuals. These differences were too easily detected, and so were differences of .0002 parts of an inch when the gauge was set close to the sizes of the plugs. With the aforementioned differences easily detected with a close set gauge, gauges with an increasingly widened gap were tried until the present size of the gap in the gauge now in use was established as the most satisfactory. Difficulty was experienced by skilled workmen, and difficulty to the same degree was experienced by boys, unskilled, and mentally retarded. It seemed that we had something here, we did not know what. The test was then extended to include ten plugs, and experimenting began.

The sense of touch..- The ability to feel small differences through the sense of touch is much greater than at first we thought it was. The difference between plugs (c) and (d) is .0002 parts of an inch and is very difficult to detect. This difference is approximately (1/37) of the gauges .0075 parts of an inch over size

when the diameter of the plug (d) is being measured. The gap of the gauge is .0075 parts of an inch larger than plug (d), and is .0073 parts of an inch larger than plug (c). Yet when this additional movement, or a lessening of the movement can be felt between the diameters of the two plugs (c) and (d) and the gap of the gauge, the plugs may be properly placed in relation to each other. Very few testees are able to feel this small difference. The differences between plugs (e) (f) and (g) is .0003 parts of an inch, the gap of the gauge at this section of the test being .0096 parts of an inch, larger than plug (g). The difference between the plugs at this section of the test in relation to the gap of the gauge is 1/32 of the gauge's .0096 parts of an inch over size. This difference can be felt with more certainty, and by more testees. The secret of this test is not only in the slight differences in diameters of the plugs, but in the use of an oversize gauge.

How much more than the sense of touch are we measuring?..- Specifically, we have learned something about the sense of touch, and we have secured enough data on the sense of touch to enable us to continue the experiment. After the test was constructed we were led to believe that in some way we were measuring more than just the sense of touch. It is possible to obtain a high score on this test in total darkness, thus eliminating the use of the sense of sight, but it has been noted that in total darkness the sense of hearing seems to supplement the sense of touch.

In addition to the sense of touch, the use of sight, and use of hearing, a fourth and more obscure sense, the sense of balance

seems to function. It is necessary in measuring the finer differences to keep the gauge level and square with the two sides of the plug while the oscillating movement used by the most successful is being carried on. This act of oscillating the gauge seems to involve a sense of balance, and any unbalanced movement which moved the gauge out of level, or out of square with the plugs would give a false impression of the differences between the plugs. All four of these senses seem to function as checks upon each other, but the sense which plays the greater part in the rearrangement of the plugs is the sense of touch. The sense of sight is secondary. The sense of hearing is supplementary, and the sense of balance seems to be called upon during measurement of the finer differences between the plugs.

A LOGICAL ANALYSIS

What may be expected of this test..- The application of one's mind to the problem seems to be involved, and those testees who can apply themselves to the task of rearranging the plugs in the proper order seem to succeed better. Maladjusted and disturbed personalities fail to score with any marked degree of success. And according to our limited experience well adjusted persons may have days when they too cannot do quite as well.

The justification:

Is this a justifiable piece of research?..- The need for the test, aside from the fact that there is none other like it, exists in the requirements of the present writer, and others, for a test which

will make possible a quick estimate of a boy's potential abilities. This test should be related to shop work and mechanical drawing, and should be a test of the keenness of the various senses used in that type of work. And if this proposed test is a good measure of intelligence and emotional stability, so much the better. For the brighter pupils verbal tests of intelligence, subject knowledge, and aptitudes are available in great numbers. For the mentally retarded pupils of the opportunity classes only a few tests are of any great value. All of the tests available for the opportunity class boys are visual if not verbal. Here, however, with the Durost-Fielding Standard Plug Test we have a test which is non-verbal, and when the testee practices the better techniques in rearranging the plugs the test becomes essentially a non-visual test with the testee depending almost entirely upon the sense of touch.

A test where four senses function..- This test opens up a new area of testing through the sense of touch, not to be dissociated, however, in this experiment, from the other senses. Here in our shops and mechanical drawing rooms every tool and instrument, and the satisfactory completion of every piece of work, depends upon the use of four of the senses, sight, touch, hearing, and a sense of balance. All of these senses are involved, to some extent, in rearranging the plugs of the Durost-Fielding Standard Plug Test. If these four senses are as essential in handling the work and the tools as they are in rearranging the plugs in the test then the relationship should be demonstrable.

The Scope of the Problem:

The construction of the test..- The full test of ten plugs was eventually constructed to the specifications previously obtained through experiences with normal individuals not connected in any way with the experimental group. It is anticipated that extensive experience with other experimental groups may make further adjustments of the plug sizes necessary, but for the present experiment it is expected that satisfactory results may be obtained.

The group to be tested..- The group for which the test was designed, our experimental group of two classes of opportunity class boys, has within it a wide variety of individuals. The range of 5 years and 7 months between the mental ages of 12 years and 5 months and 6 years and 10 months represents considerable range in ability which should be comparable to ability measured in test performance. Most of the critical data has already been presented, but the difference in mental ages is the most outstanding difference and may be the most statistically significant difference. The boys come from a wide variety of homes and their social adjustments appear to cover the whole range of social adjustments found among boys on the junior high school level. Some of the boys have high ideals of adjustment toward the present social order, while others reach a depth about as low as the human creature can go and still be out of custodial care.

What may be expected in regard to this opportunity class group..- Normal measures are not to be expected from testing experiences with opportunity class pupils. It may be, however, that

the sum total of all these boys possess in native ability will be measured more completely than the test will ever measure such ability in more highly educated and advanced individuals. The test was designed as a simple test for use in estimating the capacity of simple minds. It may be, however, that the Durost-Fielding Standard Plug Test will be found useful in estimating the capacity of normal individuals for work which may be closely related to those abilities used in rearranging the plugs of the test in their proper order. If, in the course of time, any work is found to be closely related to the abilities to perform with skill in re-arranging the plugs of the test, and that work does not involve any extensive abilities not related to performance on the test, an opportunity for work in that line will be thought to be open for the higher scoring boys of the opportunity class type. From our, at the present time, limited experience with the scores of normal individuals it is expected that many of the boys of the opportunity classes will be found with scores comparable to normal scores. Among the boys within the opportunity classes the test has already proven to be one of usefulness in quickly indicating the boy's chance of adjusting himself to the work of the shops and the mechanical drawing rooms.

CHAPTER IV

RESEARCH PROCEDURE AND TECHNIQUE

Introduction to the selected project..- The selection of a project which would interest all of the pupils is of utmost importance. Should the selected project be one which because of its simplicity would not interest the more skilled pupils, then they would not, because of lack of interest, apply themselves to the work. The experiment then would become contaminated by the more skilled pupils neglecting to do their best and the less skilled pupils excelling them because the project was within the range of their skill and interest.

What are the qualifications of an acceptable project?..- The project must interest all of the pupils and must be composed of various parts all in themselves simple enough but when assembled representing considerable skill, effort, and attractiveness. If the project has parts which can be given motion after assembly, this factor will be one of interest and will attract the attention of many pupils. If the project is one of considerable utility and will satisfy some well-felt need within the home of the pupil it will classify as a project which will bring out the pupil's best effort. If the project is composed of a number of parts, each part representing a stage of difficulty, this will allow all of the pupils to be kept busy at work which they are able to do up

to the time when all of the pupils would be graded according to the condition of their work and their rate of progress. The project shown in the accompanying photograph is such a project. A number of postulates, mostly regarding interest, are necessarily involved in the use of the single project criterion KNOWN IN THIS EXPERIMENT AS PROCEDURE A. The following photograph reveals something of the nature of the project criterion and the postulates necessarily involved will be introduced in advance of the actual research procedures.

The photograph is that of a shop project known as a wash cloth or towel hanger. A number of them may be used about the home, in the bathroom, kitchen, or pantry. In making the project the pupil learns to select stock, to saw iron and wood, to measure, to layout and drill, to file, to sandpaper, to bend steel or iron, to shave wood, to rivet, to varnish and to rub down varnish, and to assemble parts. In these activities the mechanically inclined pupil has an opportunity to show his prowess. Procedure A. revealed the relative amount of the pupil's skills.

Postulations Pertinent to the Experiment

The Control of Interest as a Variable:

Postulation No. 1 Interest in the work..- Pupils do better work on some project in which they have a personal interest rather than on a project in which, for them, interest may be entirely lacking.

Postulation No. 2 Interest in the work as a variable..- Should any project be selected and used as an objective measure of the pupil's abilities, interest or lack of interest would constitute an uncontrolled and intangible variable.

Postulation No. 3 Pupils may vary in their interests.-- Interest in regard to any arbitrarily selected project would fail to meet the requirement of interest among the members of any given group. Interest would vary from a considerable amount to an almost entire lack of any interest.

Postulation No. 4 Could more than one project be used?..- With the opportunity class pupils a multiplicity of projects, arbitrarily selected, would kill off the interest in all such arbitrarily selected projects. The regimentation necessary to control such a procedure would be intolerable to them and would result in resentment defeating the purpose of the experiment.

Postulation No. 5 One good project..- A well selected project, brought to completion in the earlier part of the year, while the initial interest in coming back to school in the fall is at its

highest point, involves about all of the interest of group magnitude that may be secured. Allowances for individual differences soon operate to cause divergences within the group, and they never again will work well upon a common project.

Therefore..- One project such as that in the accompanying photograph is all that may be used in this situation to determine whether the scores on the Durost-Fielding Standard Plug Test will indicate, with a reasonably predictive efficiency, the differences in the abilities of the pupils. Possibly some other project may be utilized in further experimentation during some other favorable period of time.

The General Hypothesis

In general pupils with the best abilities and aptitudes for the work of the shops and mechanical drawing room produce the better projects, and likewise the pupils with the poorer abilities and aptitudes produce the inferior projects This is our only working criterion of the pupil's ability and aptitude for the work.

The Fundamental Question

Will higher and lower scores on the Durost-Fielding Standard Plug Test be comparable to higher and lower performance in terms of work in the shops and mechanical drawing rooms, and can the test be used to predict such performance significantly?

The Shop Criterion

A piece of work known
as a shop project

RESEARCH PROCEDURE

The Shop Criterion

Procedure A

As a measure of the grade of work being done by the boys of the opportunity classes the boys were all given the same identical instructions in regard to the one single project which became known as the shop criterion. This project, the photograph of which accompanies the thesis, is the first one of the school year 1948-1949. Six of the parts had to be made by the boys, and all of them are simple in their construction. The project was finished by the more proficient boys in thirty-two shop periods of forty minutes each. The less proficient boys had by that time made progress, but having not completed the project had only on hand the evidence of their individual abilities and progress.

At the end of the thirty-second period the boys were told to line up at the end of the shop, each with his own project in his hands. No pupils knew in advance that such a procedure was to be followed. Each boy, having left his tools and bench, was told to look at his project and go as far up toward the head of the line as the other boys who may have had better made projects would let him. It was interesting to note the actions of some of the smaller boys who had some of the better made projects. Neither belligerence nor size prevailed. The larger boys made way for the smaller boys on the evidence of superior workmanship.

All of the boys entered into the work of judging, and it seemed as if each project was judged by at least twelve individual judges. Each judge had a personal interest in the work, had done the work himself, knew what the problems were, and considered himself a competent judge. Before all of the differences of opinion were resolved, and agreements had been reached satisfactory to all, a half an hour had elapsed. At the end of that time, it seemed to the present writer, an accurate order of rank had been established. This part of the research procedure, known as procedure A, included twenty-three boys, all that were in attendance at the school at the time. The results of this procedure when it is ranked in relation to the boys' ability to score on the Durost-Fielding Standard Plug Test, produced a correlation (rho) of .41. When the problem and the following research procedures are considered as a whole in relation to the testing instrument, such a correlation must be considered as significant.

Procedure B:

All of the boys' mechanical drawing papers were ranked in order by the present writer. This is a procedure which has been carried on four times a year for the last twelve years and is a familiar and practiced one. The drawings were sorted in such a way that the boys' names were not in evidence, and this too is a long-practiced procedure. The names were compiled after the sorting. A rank order thus established was then compared to the rank order of the boys' scores on the Durost-Fielding Standard Plug

Test, and a resulting correlation (rho) of .46 was obtained. There were twenty-eight boys involved in this experiment. All were opportunity class boys, but much of their work is comparable to the work of seventh and eighth grade boys. Considering the relationship between mechanical drawing and the need of such knowledge where work in the shop is concerned, this correlation of .46 is considered to be of some substantial significance.

Procedure C:

A rank order of the boys' intelligence quotients was derived from the regular school records and psychological reports. The Binet-Simon intelligence ratings, in the case of these boys, do not vary greatly year to year, but the last recorded intelligence quotient was taken in each case. The correlation resulting from a comparison of the two rank orders, that of the Binet-Simon intelligence quotients and the Durost-Fielding Standard Plug Test, was obtained as of .33. This correlation is considered to be only a slight correlation, and of very little significance. However, further experimentation with the Binet-Simon mental age rating revealed that chronological age, a component of the intelligence quotient, operated to deflate the relationship between the boys' intelligence quotient and the boys' scores on the Durost-Fielding Standard Plug Test. So further procedure follows in the direction of the boys' mental age as a factor separate from the intelligence quotient.

Procedure D:

In this procedure we find that the correlation (rho) between the scores on the Durost-Fielding Standard Plug Test and the chronological ages of the twenty-eight boys of this experiment to be .01. The three years and six months difference between the lower range of twelve years, and the upper range of fifteen years and six months make no difference at all in the boys' ability to score on the test.

Procedure E:

Mental age is now a factor in which we have become concerned because of its substantial correlation (rho) of .56 with the Durost-Fielding Standard Plug Test. The range of mental ages from six years and ten months to twelve years and five months, a range of five years and seven months, does represent a considerable difference in ability to score on the Durost-Fielding Standard Plug Test. From these indications, and in terms of this low level of mental ages, it is possible to say that the relationship between the Durost-Fielding Standard Plug Test and the Binet-Simon mental age is one of substantial and marked relationship which cannot be ignored.

Procedure F:

There are many variables involved in estimating a boy's capacity for shop work which cannot be measured through our usual objective techniques of testing.

For individuals who cannot estimate a boy's capacity..- And while considering the many variables involved and uncontrolled it may be just as well to understand that for some individuals it is not at all possible for them to estimate a boy's capacity for work in the shop with any degree of proficiency. Objective testing and statistical methods are a valuable adjunct to the judgments which necessarily have to be made by such an individual.

Considering the individual who can estimate a boy's capacity for work in the shop..- On the other hand, there may be some individuals who possess the ability to estimate a boy's capacity for work in the shop with a proficiency that for practical purposes can suitably be applied to most normal situations.

Selecting personnel without objective testing..- The present writer has been selecting personnel to carry on many varied types of work to a reasonably successful conclusion for a period of over thirty years. This was done without the aid of objective testing. No poor results have ever been in evidence, no outstanding failures have occurred as evidence that the present writer could not select personnel capable of doing correctly what they were told to do. However, the present writer never before has had an opportunity to measure objectively this so-called faculty. With considerable misgiving, and with much trepidation, an attempt will be made here to gain some statistical evidence in regard to what degree, if any, the present writer does possess such a faculty for estimating the capacity of other individuals for work of a general nature in shops.

[illegible]

[illegible]

[illegible]

[illegible]

The statistical evidence..- The quarterly marks for the first quarter of the school year, and the rank order which is the basis of these quarterly marks, were compiled in consideration of not only notable production of the school shop projects, but in consideration of such variables as workman-like attitudes, understanding of what one may be doing, interest in the work, and cooperation in matters of shop routine. The compilation of the results of the testing with the *Durost-Fielding Standard Plug Test* was delayed until the rank order concerning the quarterly marks was complete. The chance that any halo effect which could operate to inflate the correlation between the *Durost-Fielding Standard Plug Test* and the compilation of the rank order from which the quarterly marks was derived was in this way controlled. The resulting correlation of .80 between the *Durost-Fielding Standard Plug Test*, and the rank order upon which the quarterly marks was based has a predictive value of 40 per cent better than chance alone. While this correlation of .80 denotes a high relationship, the predictive efficiency of 40 per cent better than chance alone leaves plenty of opportunity for failure. With a predictive value of only 40 per cent better than chance as a prevailing hazard in the work of the present writer, and having given chance a period of over thirty years in which to operate, it is a wonder that some one hasn't been killed. But perhaps the present writer and his associates are more careful when the chips are really down.

Procedure G:

THE RELIABILITY OF THE TEST

The test and re-test technique.-- There is to some extent experimental evidence available that the Durost-Fielding Standard Plug Test possesses a reasonable degree of reliability. In considering the reliability of this test, however, it must be considered that the test may be more reliable than the group being tested. A correlation between the test rank order as first obtained, and a rank order obtained through testing the same individuals three months later produced a correlation of .62. If two of the cases of the thirty-two cases involved in this technique, which accounts for a difference in rank resulting in 40 per cent of the total of the sum of the differences squared were thrown out, the resulting correlation would be enhanced to the extent of .73. However, the chance was taken with the original thirty-two cases who are of a group of individuals many of whom are considered to be quite unstable, so in relation to this experimental group the reliability of the test must be taken as of .62.

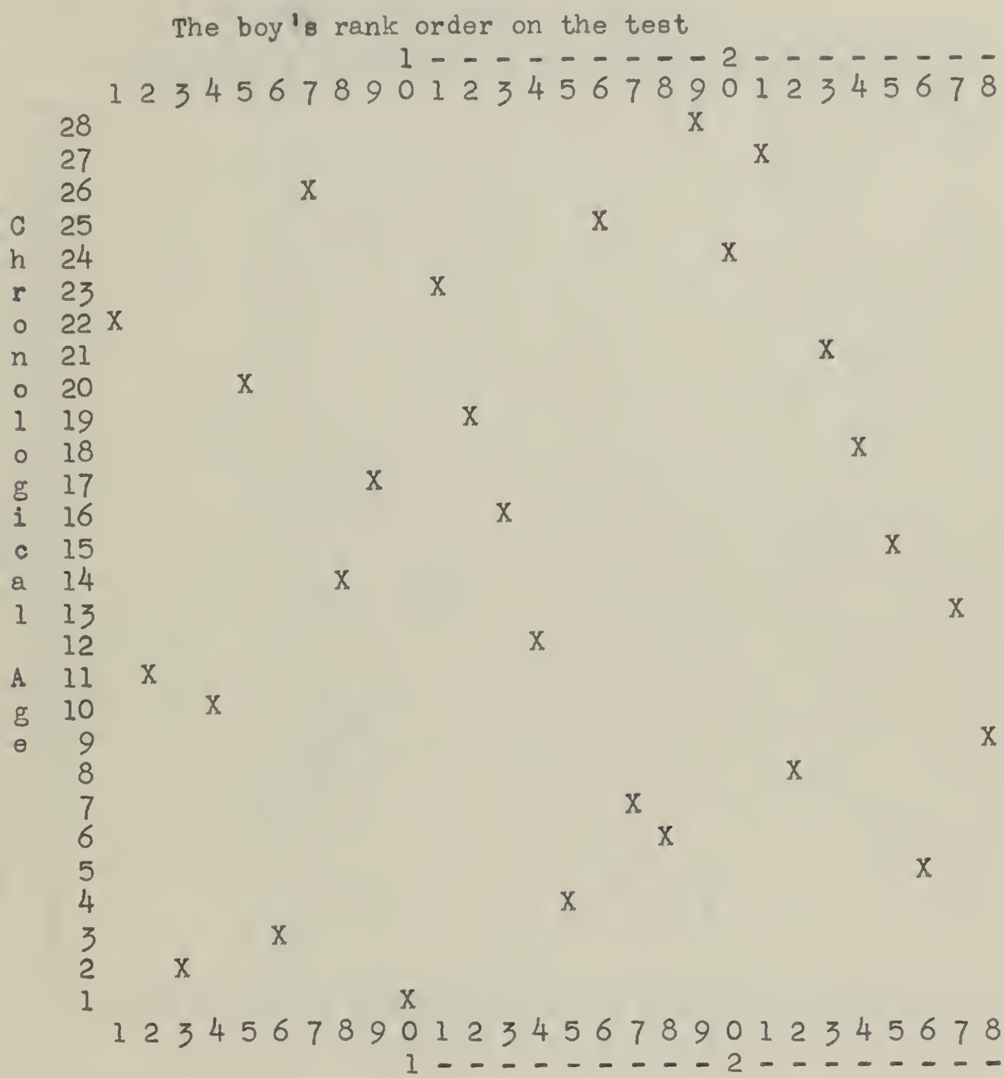
Procedure H:

An experiment in comparing the performance of the opportunity class pupils to the performance of ten superior students working on the graduate school level in Dr. W. N. Durost's seminar in measurement at Boston University, Jan. 1949.

Durost-Fielding Scores	Opportunity Class Scores	Graduate Student Scores
2	X X	X
4	X	X
5		
6	X X	X X
7	X X	
8	X X X	X
9		----------Mdn. 9
10	X X X X X	X X
11	X	X
12	X X	
13		
14	X X X X X---Mdn. 14.1	
15		
16	X X	X
17	X	X
18	X X	
19		
20	X X	
21	X	
22	X X	
23		
24		
25		
26		
27		
28	X	
29		
30	X	
31	X	
32	X	

The actual difference between these two medians is 5.1 with a probable error of plus or minus 1.434. The difference between the medians is sufficiently large to insure a true difference greater than zero in 99.0 cases in 100 in favor of the graduate student group.

A scatter plot showing in diagram the correlation (rho) of .01 of the <u>Durost-Fielding Standard Plug Test</u> against the rank order of the boy's chronological age.



Comment:

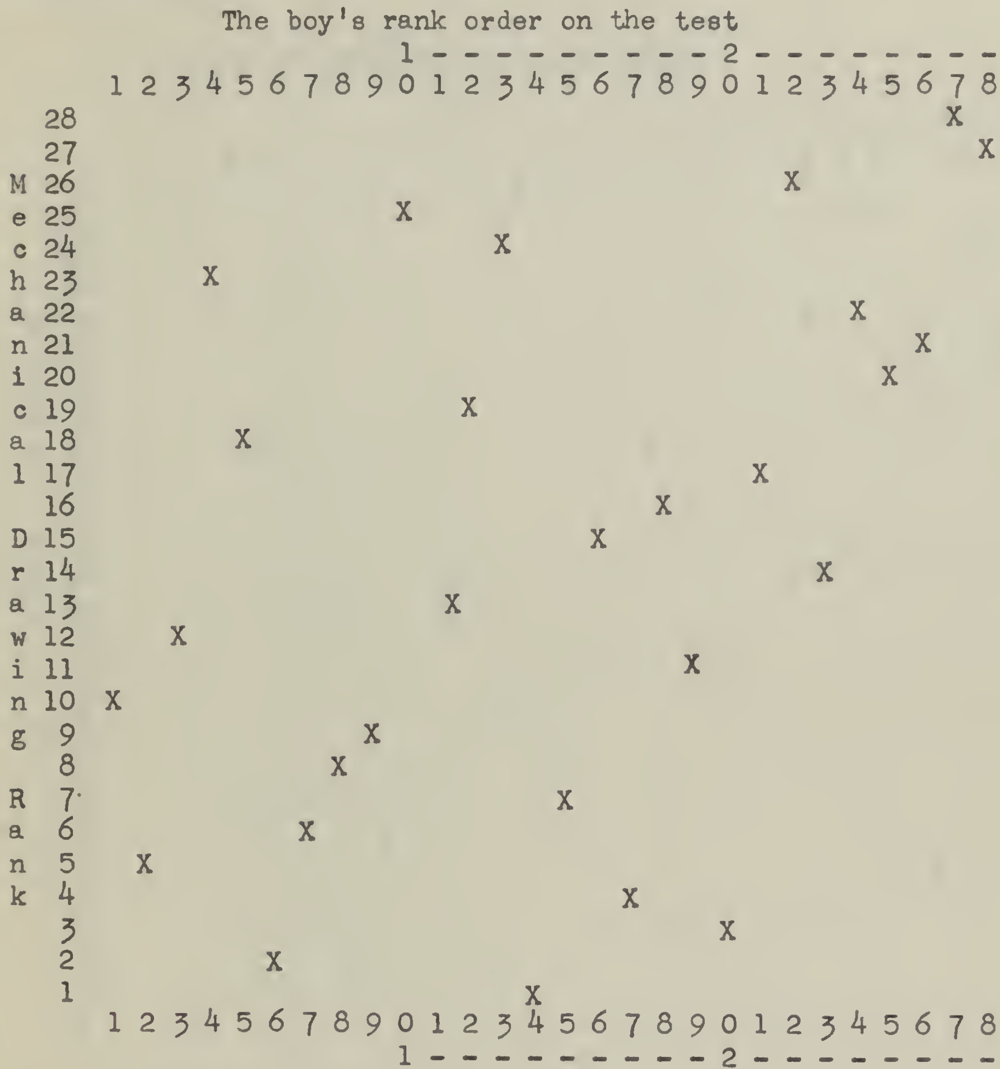
The range of the boys' chronological ages from 12 years to 15 years and 6 months has no effect upon the experiment in regard to ability to score in relation to age.

The diagram shows no correlation, and age may be ignored in its relationship to this experiment.

[illegible]

[illegible]

A scatter plot showing in diagram the correlation (rho) of .46 of the Durost-Fielding Standard Plug Test against the boy's rank order in mechanical drawing.



Comment:

While a correlation of from .35 or .40 to .50 or .60 is considered to indicate a "markedly present" relationship the scatter plot would indicate that this relationship is nothing in which to place any great dependence.

A scatter plot showing in diagram the correlation (rho) of .56 of the Durost-Fielding Standard Plug Test against the boy's rank order in mental age (Binet)

```
        The boy's rank order on the test

                       1 - - - - - - - - - 2 - - - - - - - -
     1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8
  28                         X
  27                                               X
  26                                                     X
  25                                     X
M 24                                           X
e 23                                                       X
n 22                                                 X
t 21                                         X
a 20                                 X
l 19 X
  18                                       X
A 17                                             X
g 16                     X
e 15             X
  14                               X
  13   X
  12                       X
  11                           X
  10                   X
   9       X
   8                 X
   7                                                   X
   6         X
   5                                   X
   4               X
   3           X
   2     X
   1                             X
     1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8
                       1 - - - - - - - - - 2 - - - - - - - -
```

Comment:

While a correlation of from .35 or .40 to .50 or .60 is considered to indicate a "markedly present" relationship the scatter plot indicates that a correlation of .56 represents a more markedly present relationship than the correlation of .46 in the scatter plot of the test against mechanical drawing.

A scatter plot showing in diagram the correlation (rho) of .41 of the Durost-Fielding Standard Plug Test against the rank order of the boy's performance on the "piece of work" known as the objective criterion or shop project

The boy's rank order on the test

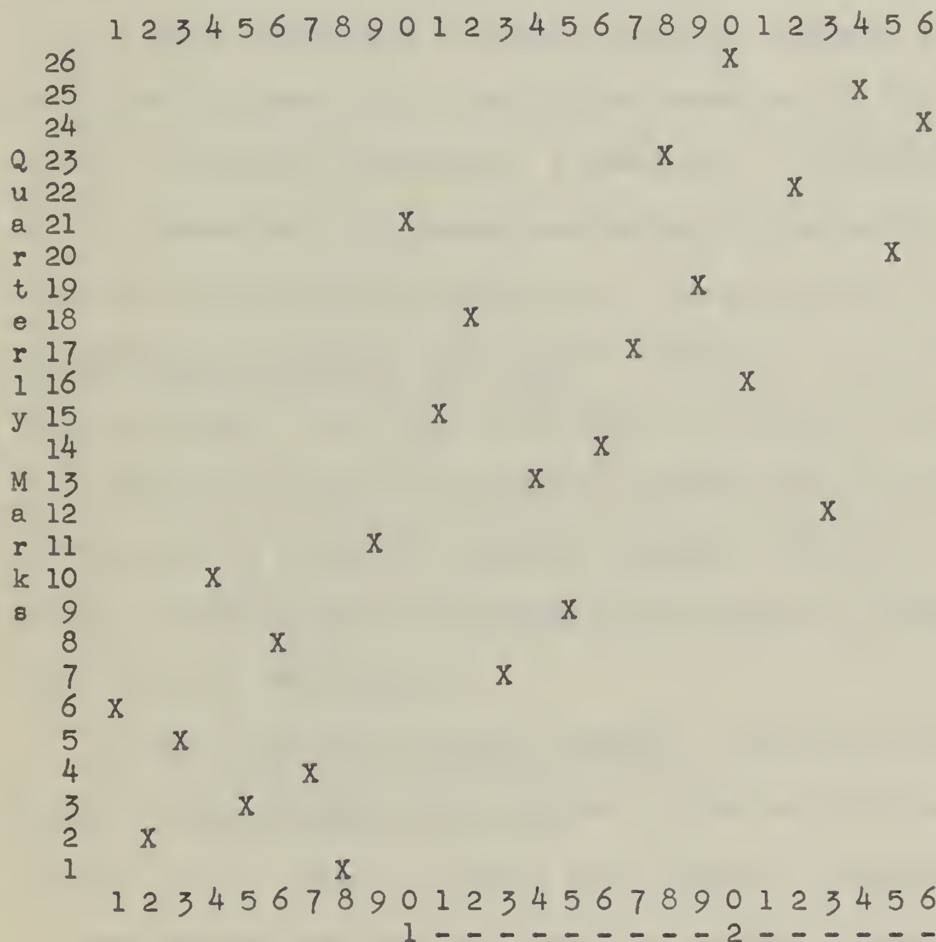
```
                       1 - - - - - - - - - 2 - - -
     1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3
  23                                           X
S 22     X
h 21                                             X
o 20                               X
p 19                                       X
  18                                 X
P 17                     X
r 16 X
o 15                         X
j 14                           X
e 13                                     X
c 12                   X
t 11                       X
  10                             X
   9                         X
   8                                         X
   7                           X
   6       X
   5                                   X
   4           X
   3             X
   2   X
   1         X
     1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3
                       1 - - - - - - - - - 2 - - -
```

Comment:

A review of cases would reveal that the Durost-Fielding Standard Plug Test IS A BETTER INDICATION OF WHAT A BOY IS CAPABLE OF than is indicated by the performance of those whose test scores are of a high order and whose work was, in this experiment, very poor. Too many variables were uncontrolled here. The test, we are quite certain, is considerably better than the correlation of .41 indicates.

A scatter plot showing in diagram the correlation (rho) of .80 of the Durost-Fielding Standard Plug Test against the boy's rank order obtained from the boy's first quarterly marks in the general shop

The boy's rank order on the test



Comment:

A correlation of .60 to .70 is considered to be high. A correlation of .80 is significantly high, and may be taken to indicate that a relationship exists to such an extent that the relationship cannot be ignored.

CONCLUSIONS

Where the test is found to be useful..- The Durost-Fielding Standard Plug Test being the only test measuring in this particular area of mechanical ability may become a specific test for any kind of work where sensitivity of touch is used in obtaining great precision in machine parts, watch-like mechanisms, and fine electrical and optical instruments. A great deal of experimenting will be needed and only through such experimenting is it possible to arrive at any real conclusions as to the efficiency of the Durost-Fielding Standard Plug Test in relation to any specific testing problem. In so far as the test is concerned at present the indications are that the test will measure with reasonable proficiency the presence of general mechanical ability in mentally retarded boys and indicate with reasonable prediction their relative mental ages.

In the direction of further research..- While the present Durost-Fielding Standard Plug Test has proved adequate for its present use in a school situation among mentally retarded boys, for more serious work the test should be extended. Perhaps this extension should be only in difficulty rather than in the number of plugs in the test. As a general test the Durost-Fielding Standard Plug Test allowed most all of those tested to experience some degree of success. For some of the specific jobs the present writer has in mind more of the plugs would

need to be much closer together in regard to size. None of the plugs would be more than .0003 parts of an inch different from the next larger or smaller plug in the series. The gauge size would remain the same. A test so constructed probably would not measure general mechanical ability at all, but just some specific and highly specialized degree of the sense of touch. Experimentation could prove this either to be right or wrong. Without experimental science we do not know that such a high degree of sensitivity of touch is not more closely related to playing a violin than to mechanical ability.

The present writer-experimenter thinks that sensitivity of touch is related to high intelligence. On the levels of lower intelligence, the higher intelligence scores are related to the higher scores on the Durost-Fielding Standard Plug Test. Usually the validity of a test is inferred by correlating it with other tests whose standards of validity have been accepted. The correlation (rho) of .56 between mental age (Binet) and the scores on the Durost-Fielding Standard Plug Test indicate that some of the factors which enable one to score on an intelligence test are also present in those abilities which enable one to score on the Durost-Fielding Standard Plug Test. While some tests of intelligence correlate with each other to a higher degree than our .56, according to Ernest W. Tiegs,[1] there are also tests of intelligence which correlate with each other to a lower degree.

1/ Ernest W. Tiegs, Tests and Measurements for Teachers. Houghton Mifflin Co., New York, 1931. p. 304.

Our correlation of .56 (rho) is high enough to be interesting.

Are we interested in how well the Durost-Fielding Standard Plug Test correlates with other tests of mechanical ability?..- In so far as how well the Durost-Fielding Standard Plug Test correlates with other tests of mechanical ability, except for one or two reasons, the present writer is not greatly interested. A test which measures something quite different should not correlate significantly with a test which does not measure in the same area. It is possible, however, that the Durost-Fielding Standard Plug Test does measure something general to all tests of mechanical ability. A (G) factor, for instance. In this case we would be getting quite close to the seat of human abilities in measuring the quality of the nerves of the nervous system and their brain connections with the sense of touch, sight, hearing, and balance. The field of experimental science is wide open here in this regard.

Is the present Durost-Fielding Standard Plug Test the final answer?..- The test as it is now constructed is suitable for the purposes of laboratory testing. It is large enough. It should not get lost easily nor may it be carried away inconspicuously. The test costing several hundred dollars to reproduce makes this a factor to be considered. The present writer, however, would like a test which could be carried inconspicuously in a side pocket of a sack coat (the ordinary business suit). Testing could then be done during propitious moments without attracting undesirable attention. Such a test would have to be completely re-designed and standardized. However, the statistical data now made available

in this experiment would definitely be of considerable importance in the construction of any similar test. With such a test available as good an estimate of a boy's possibility of adjustment to shop work and mechanical drawing may be obtained within ten minutes as may be obtained from several weeks of school experience with the boy.

Factors of safety and the Durost-Fielding Standard Plug Test..- In the present writer's work it is well to discover immediately the status of the boy's intelligence. With a test score as easily obtained as that of the Durost-Fielding Standard Plug Test score, and the Sims score, both the boy's intelligence and his mechanical aptitude may be estimated. This is important. Psychological reports are not always immediately forthcoming, and during the period of administrative lag unstandardized techniques are used to discover whether a capable boy or a borderline idiot is being presented for admission to the class. The risk is considerable. According to Max A. Henig,[1/] in his study of the relationship of intelligence to the percentage of accidents among his group of one hundred and sixty-four boys, those boys with class (D) intelligence were given a hundred per cent accident rating, (C-) seventy-one per cent, (C) sixty-four per cent, (C+) fifty-one per cent, (B) thirty-two per cent, (A) thirty-three per cent. How we avoid accidents among opportunity class pupils is an art in itself. In this one hundred per cent area of low intelligence and accident propensity an investigation wouldn't reveal enough

1/ Max A. Henig, "Intelligence and Safety", Journal of Educational Research, March, 1929.

data to make a statistically significant report. However, we know that the accident-prone individual is one hundred per cent present among the boys of the opportunity class. Max Henig's study is an old and familiar one here. With the Durost-Fielding Standard Plug Test which measures both mechanical ability and intelligence the possibility of an interesting study of the relationship of accident proneness to this test performance is possible.

Is the Durost-Fielding Standard Plug Test difficult enough for normal groups?..- The test was constructed as a non-verbal test for use among sub-normal groups and proves very satisfactory for that purpose. The median score for our sub-normal groups was a score of 14.1 where theoretically it should be 16.5. Very little need for correction is indicated and none is contemplated. Of ten superior individuals working on the graduate school level, the median score was 9. While the test appeared easier for them, none of them obtained a perfect score. As indicated also by the sub-normal group, degrees of higher intelligence are indicated through superior performance with the Durost-Fielding Standard Plug Test apparatus.

This test, which was designed to allow some very sub-normal individuals a chance to score above a score of 33,— a score which in this test represents the mean chance deviation — achieved that purpose. Scores below 33 being due to chance alone have no prognostic values. We avoided them. Perhaps we have hit it right this time. Not speculation but experimental evidence

must be our guide. On the basis of the experimental evidence available as a result of this tentative exploration we can come to no more extensive conclusion. More experimental work should be done before this test is altered to meet some theoretically statistical criterion. The changing of some of the less critical plugs would not entail any great difficulty.

Is it practical from the viewpoint of manufacturing to reproduce the Durost-Fielding Standard Plug Test?..- The present test apparatus could not be reproduced with any exactness for much less than one thousand dollars. Our method of selecting the plug sizes from the random turnings of the boys of our opportunity classes, while good experimental procedure has led us into difficulties. This method, however, may now be abandoned. We know now what is needed. Should the letter sizes (T).3580", (U).3680", and (V).3770" be used in the choice of diameters for the selection of drill rod as the basis of the plug sizes the labor in manufacture would be minimized. Each drill rod size need be ring lapped but a few tenths, or at the most one thousandth of an inch under the manufacturer's size for the rod. If the plugs were hardened under carefully controlled conditions the ring lapping would become much less expensive. Automatic screw machines could be used to produce the rough plug up to the point of hardening and ring lapping. The next test apparatus to be made by the present writer will take into consideration the conditions to be met in the problems of reproduction. Producing the gauge is also considered. The problem of producing the gauge

by hand would prove to be an expensive one. On a punch press and stamping proposition the costs would be cut tremendously. The gap of the gauge could be ground to reasonable tolerances in a gang fixture, thus eliminating the costs of expensive lapping. If the knurled handle should be abandoned in any labor-saving process of manufacture some loss would accrue in psychological rapport. There are sound reasons for such a handle in the present design of the plugs. The knurled handle and necked-in section is standard practice in the manufacture of commercial test plugs. To the mechanic this particular design is a familiar one. Therefore, to redesign this very practical mechanic's tool and then try to secure the co-operation of skilled mechanics in any future experiment would result in erecting a psychological barrier between the mechanic and the test administrator. The presence of a familiar tool between the test administrator and the mechanic would tend to insure better co-operation. The present test may be a better test in its present form without any of the changes about which we have speculated, and in the absence of any experimental evidence to the contrary we had better let the matter rest until more experimental evidence is produced.

In view of the correlations obtained in experimenting with this very unstable group — and under the uncontrollable conditions of their abnormal lives — one should find some indication that a valuable test has been made available. We do not think that the test need be limited to work with sub-normals.

Considering the sense for rhythmic vibrations and tones..- In the selection of the plugs from the random sizes which had been turned, gaps were left so that there would be no orderly graduation of sizes from the largest plug to the smallest. These irregular gaps in the distribution of the plug sizes were supposed to prevent the operation of any keenness of the sense of rhythm which might enable some testees to rearrange the plugs of the apparatus through that sense. We did not know at the time that other senses beside that of the sense of touch would be factors in the testee's performance. It was thought that rhythm was a sense of more importance to the musician than to the mechanic. As a result, an attempt was made to interrupt any progressive and orderly sounds and feeling that progressive increments of distance and tone might engender. It is thought now that this was a great mistake.

Rhythmic vibration and degrees of tone and the mechanic..- An intelligent mechanic is conscious of such rhythms in his work. Such rhythms and tones of sound as the mechanic knows them are for the most part the results of the forces applied to his work. Such sounds and vibrations are usually steady and rhythmic. When such sounds and tones and vibrations either increase or decrease in intensity the mechanic, alert for such changes, knows what action to take.

It was thought that, through training, one who was not a mechanic but a skilled musician perhaps could achieve a high score through a trained sense of tone and rhythm. For this reason

alone any possible rhythmic progression was broken.

Rhythm and tone - the musician and the mechanic..- The rhythm of the Blue Danube Waltz was set to the rhythmic changes of the paddle wheels of the side-wheel steamboats of the river. Here in the ever-changing forces of the stream the struggle of the boat against the stream set up a rhythm. The captain felt it, the engineer felt it, and the judgment of both was guided by it. To the captain the changes in rhythm meant shoals and bars, deep water, and force of the current. To the engineer the changes revealed the power of his engine and its condition. To Johann Strauss the vibrations became the rhythm for a waltz that does not die. The difference between the great musician and the great engineer is probably only one of culture.

Our mechanics and engineers understand and live by the rhythmic pulse of their engines and the roar and hum of their prime movers and machine tools. Our trying to avoid a measuring of this sense of tone and rhythm was a mistake. To figure out the most efficient series of test plugs based upon some natural and infinitesimal increment in degrees of sensitivity of touch, tone, rhythm, sight, and balance would constitute a considerable contribution to the new science of cybernetics. There is much that we do not yet know, so again the matter must await a further advance of knowledge in this direction. WE BELIEVE, HOWEVER, THAT AS A RESULT OF THIS CONTRIBUTION TO KNOWLEDGE THE DREAM OF A MASTER TEST IS MUCH BRIGHTER.

BIBLIOGRAPHY

Baker, Harry J., and Crockett, A.C., Detroit Mechanical Aptitudes Examination for Boys, Public School Publishing Co., Bloomington, Illinois.

Bennett, George K., Mechanical Comprehension Test, The Psychological Corporation, 522 Fifth Avenue, New York, N.Y.

Bennett, George K., and Cruikshank, Ruth M., A Summary of Manual and Mechanical Ability Tests, (Preliminary Form), The Psychological Corporation, 522 Fifth Avenue. New York, N.Y.

Berkshire, J.R., and others. "Test Preferences in Guidance Centers", Occupations 3'48 p. 339.

Bingham, Walter van Dyke, Aptitudes and Aptitude Testing, Harper & Brothers Publishers, New York, 1937. 390 p.

Conrad, Herbert S., "The Scoring of the Rearrangement Test", Journal of Educational Psychology, 27: Apr. '36. p. 241-252.

Crawford, J.E., and Crawford, D.M., Crawford Tridimensional Structural Visualization Test, The Psychological Corporation, 522 Fifth Avenue, New York, N.Y.

Curston, E.E., and Dunlap, J.W., "Scoring the Rearrangement or Continuity Test", The School Review 1930 38: p. 613-616.

Dunlap, Jack W., and Kurtz, Albert, Handbook of Statistical Nomographs, Tables, and Formulas, World Book Company, Yonkers on Hudson, New York, 1932. 163 p.

Henig, Max A., "Intelligence and Safety", Journal of Educational Research (Sept. 1927). p. 82

MacQuarrie, T.W., The MacQuarrie Test of Mechanical Ability, The California Test Bureau, 3636 Beverly Boulevard, Los Angeles, California, 1925.

O'Connor, Johnson, Wiggly Block, C.H.Stoelting Co., 424 N. Homan Ave., Chicago, Illinois.

O'Connor, Johnson, O'Connor Finger Dexterity Test, (Worksample No. 16.), Human Engineering Laboratory, Stevens Institute, Hoboken, New Jersey.

O'Rourke, L.J., O'Rourke Mechanical Aptitude Test- Junior Grade, The Psychological Institute, 3506 Patterson St., N.W., Washington, D.C.

Patterson, Donald G., and others, Minnesota Mechanical Ability Tests, The University of Minnesota Press, Minneapolis, Minnesota, 1930. 586 p.

Sims, Verner M., "The Scoring of the Rearrangement Test", Journal of Educational Psychology, 28: 1937 p. 302-304.

Stenquist, J.L., Assembling Test of General Mechanical Ability, C.H.Stoelting Co., 424 N. Homan Avenue, Chicago, Illinois.

Tiegs, Ernest W., Tests and Measurements for Teachers, Houghton Mifflin Co., New York, 1931 p. 304, 470 p.

Tiffin, Joseph, Industrial Psychology, Second Edition, Prentice Hall Inc., New York, 1947. p. 126-7, 141.

Worcester, D.A., "Still Further Comments of the Scoring of Continuity Tests", The School Review 1930 38: p. 462-466.

Ziegler, W.A., Minnesota Manual Dexterity Test: Minnesota Rate of Manipulation Test, Publisher, Educational Test Bureau, Minneapolis, Minnesota, 1931.