

U. M.

THE UNIVERSITY OF MINNESOTA

GRADUATE SCHOOL

Report  
of  
Committee on Thesis

The undersigned, acting as a Committee of the Graduate School, have read the accompanying thesis submitted by Mery Grace Arthur for the degree of Master of Arts.

They approve it as a thesis meeting the requirements of the Graduate School of the University of Minnesota, and recommend that it be accepted in partial fulfillment of the requirements for the degree of Master of Arts.

Herbert Woodrow  
Chairman

Rupert C. Lodge

M. J. Van Wagenen

August, 1918

UNIVERSITY OF  
MINNESOTA  
LIBRARY

THE UNIVERSITY OF MINNESOTA

GRADUATE SCHOOL

Report

of

Committee on Examination

This is to certify that we the undersigned, as a committee of the Graduate School, have given Mary Grace Arthur final oral examination for the degree of Master of Arts . We recommend that the degree of Master of Arts be conferred upon the candidate.

Minneapolis, Minnesota

June 26 1918

Herbert Woodhew  
Chairman

R. C. Lodge

M. J. Panwagener

253572

1918 15 13.0. 98

An Absolute Intelligence Scale  
A Study in Method

A Thesis submitted to the  
Faculty of the Graduate School of the  
University of Minnesota

by

Mary Grace Arthur

In partial fulfillment of the requirements  
for the degree of  
Master of Arts

August

1918

NOM  
7A:7

## An Absolute Intelligence Scale

### A Study in Method<sup>1</sup>

#### Chapter I

#### Introduction

The present investigation is an attempt to arrange a series of group tests which can be used:

- (1) to classify children of grade school age according to mental ability,
- (2) to test the possibility of constructing a scale of absolute intelligence units,
- (3) to form the basis of a comparison of the "more or less" with the "all or none" method of scoring,
- (4) in a system of weighting which is neither arbitrary nor determined by subjective evaluation of the tests used, but which is logical in its origin and automatic in its application.

<sup>1</sup>This work was done under the direction of Dr. H. Woodrow of the Psychology Department of the University of Minnesota.

It is with pleasure that acknowledgment is made also to Miss Frances Lowell, of the Research Department of the Minnesota State School for Feeble-Minded for suggestions as to the choice of tests and details of procedure; to Miss Edith Taylor, principal of Hancock School (the school tested) St. Paul, Minn. for her aid in securing a favorable attitude towards the tests on the part of teachers and subjects; to Mr. B. B. Jackson, Superintendent of Schools of Minneapolis, Minn.; to Mr. S. O. Severson, principal of Simmons School (Minneapolis) where the trial tests were given; and to the teachers of both schools where testing was done.

### Choice of Tests

As the tests were to be given to all the pupils of a grade school, it was necessary to select such as would be simple enough for the younger children to comprehend, and at the same time would be difficult enough to tax the ability of the most mature.

Those tests were chosen which were as diverse as possible in the nature of the functions they were supposed to involve. Of several tests apparently involving the same functions, that one was selected which, from the work of previous investigators, seemed to show the greatest increase in ability from one age group to the next. The tests finally decided upon were:

Immediate Memory

Easy Opposites

Hard Opposites

Substitution

Word Building

Language Completion

Anagrams

Cancellation

Comprehension

### Subjects

The subjects were grade school pupils from 6 to 17 years of age, inclusive. Norms for the separate tests were secured for the ages of 6 to 13 years, but no attempt was made to classify the six year old subjects. Results for the age groups above 13 were discarded as all the members of these groups were retarded peda-

gogically. Eleven subjects of the 13 year group were one half to one year advanced pedagogically, and entered high school before the testing was completed.

For all subjects excepting those in this small eighth grade group, chronological ages were reckoned to March 8, as that date would give an error of not over one month for the majority of the cases. Ages were grouped according to the popular usage, i. e., 6 years included subjects from just 6 years to 6 years, 11 months.

The 673 subjects whose scores were retained include practically every grade of intelligence between definite feeble-mindedness and genius. The known range based on individual tests, given according to the Kuhlmann 1917 Revision<sup>1</sup> of the Binet-Simon scale, is from an I. Q. of 75 to one of 134.

The upper and lower middle classes are represented largely in the district in which the school tested is situated. None of the subjects come from the families of the very rich; few are very poor.

#### Time

Most of the testing was done between February 8, 1918 and April 8, 1918, but about 50 stragglers were tested between April 8 and May 8. An eighth grade group was given the first six tests in the early part of January of the same year.

The testing was done at all hours from 8 A. M. to 4 P. M. but the best hours of the day were given to the younger groups. To test each group required from two and one half to

<sup>1</sup>Kuhlmann, 1917, The Measurement of Mental Development.

three hours, but no group was allowed to work more than ninety minutes at a time even when exercise between tests was allowed.

### Conditions

Conditions, both objective and subjective, were kept as nearly uniform as possible for all groups. Weather conditions were disregarded excepting one day when they were particularly depressing and a number of slow groups, to whom the experimenter was not familiar, were being tested. On this day an unusual amount of play was allowed between tests.

Restless groups made up of younger subjects and one sixth group which fatigued quickly, were given the opportunity to exercise and relax at one half to one hour intervals. Although objective conditions in this respect were not uniform, the subjective conditions were probably more so than if a fixed order of procedure for all age groups had been adhered to. Praise and encouragement were given in fairly definite amounts.

With rare exceptions, the subjects seemed to find real enjoyment in the tests. Without urging on either the part of teacher or experimenter they came early, stayed after school and came to be tested on half days when they were not required to attend school. Nothing was told them of the object, purpose or possible results of the testing.

A fixed order for the tests was impossible, owing to the necessity of taking groups when they could be had and of using occasional half hour intervals, but a general plan was adhered to when practicable.

Teachers, though present, remained in the background

excepting in the lower grades where they helped in the distribution of materials and in the filling out of the Trabue practice sheets, etc.

### Instructions

Written instructions were used by the experimenter in all tests and were half spoken, half read. Most of the instructions used were adapted from those of earlier investigators. They were tried with first, second, sixth and eighth grade groups of another school, and were then revised on the basis of the results with these groups. These revised forms were tried on the eighth grade group which was about to enter high school from the school which was to be tested. As the results from this group seemed to be satisfactory, the instructions were retained without change for the other groups, and the scores obtained from this group were included in the final averages.

In giving instructions, group response in a trial test was secured wherever possible. Purity of diction was sacrificed in every case where a colloquial form appeared to be more readily understood.

### Scoring

The test papers, with the exception of some of the cancellation and comprehension sheets, were all scored by the experimenter, and these were done under her immediate supervision. The hard opposites, language completion and most of the comprehension papers were gone over a second time either on a different day from that of the first scoring or by a different person. This was not



only to insure against mechanical errors but to avoid the results of a possible shifting of standards during the progress of the scoring. For the other tests, papers selected at random were re-scored, but so few errors were found that for the rest of the papers the original scoring was assumed to be accurate.

## Chapter II

## Description of Tests

Digits

With every group, the test for immediate memory was the first test given. Two series of digits were given in immediate succession with auditory presentation and written response. Both series were taken from the Yerkes-Bridges<sup>1</sup> point scale and were extended downwards to include groups of two and one digits. The two-digits groups were taken from the three year test of the Kuhlmann<sup>2</sup> (1917) Revision.

In the first grade rooms the teacher was asked to write the digits 1 to 9 on the blackboard and to have the children name them as she wrote. This precaution is absolutely necessary if the test is to have any value as a memory test for children of this grade. In almost every group some child had to be reprimanded for attempting to write during the reading of the digits, but after this had occurred the subjects seemed to understand that all instructions were to be followed exactly.

In scoring responses the method used in other minimal change experiments was used. A single displacement of digits in a group was scored as half an error. If more than one displacement occurred the group was counted as all wrong. The average of the two series was taken and recorded as the final score.

The test was not regarded as being especially valuable but was included because of its wide use in other scales and for

<sup>1</sup>Yerkes, Bridges and Hardwick, 1915, A Point Scale for Measuring Mentality.

<sup>2</sup> Kuhlmann, 1917, The Measurement of Mental Development.

the purpose of comparing it with other tests.

### Opposites

The easy and hard opposites tests were chosen because of their believed high correlation with general intelligence. The two lists were selected from the frequency tables of King and Gold.<sup>1</sup> The easy opposites have an average frequency of 99 for educated adults. The hard opposites have an average frequency of about 93. The advantage of basing the lists upon frequency tables instead of upon association time is the greater ease of scoring responses. In the easy list only one word has more than a single correct response, and only one alternative exists even there. Seven words of the hard list have more than one exact opposite. In scoring the responses for these, the standards of King and Gold were adhered to strictly excepting that no half credits were allowed and that the word "hatred" was accepted as the opposite of "love".

In scoring, spelling was disregarded as much as possible. Any word that was phonetically correct or nearly so was given full credit unless it had a definite meaning other than the correct one. For instance, "nēro" was accepted as the opposite of "broad", but "know" as the opposite of "yes" was regarded as an error. No changes of tense were permitted, as "awoke" for the opposite of "asleep", nor were approximate opposites, as "cool" instead of "cold" as the opposites of "hot" given any credit.<sup>2</sup>

The hard opposites were not given to any grades below the 2A, nor the easy opposites below the 2B, as the younger sub-

<sup>1</sup>Journal of Educational Psychology, 1916, p. 459.

<sup>2</sup>See Appendix for lists of opposites with adult frequencies and accepted responses.

jects had not had school training enough to enable them either to read the stimulus word or to write the response.

Thirty seconds were allowed for the easy list for grades 4A to 8A inclusive, and sixty seconds for all grades below the 4A. Scores from the latter groups were divided by 2, in order to make them comparable with those of the higher groups. For the hard opposites, sixty seconds were allowed in all grades.

### Substitution

For a substitution test, the Woodworth and Wells form was selected as it appeared to be better suited to the ability of younger subjects than are some of the other forms that have been used in previous investigations.

### Word Building

Just before the word building test, special praise and encouragement were given to the older groups and a few minutes of relaxation to younger groups, as this list appears to make heavier demands upon attention and to require more sustained effort than any other of those used.

The two forms, e a i r l p and a e o b m t for which norms were given by Whipple<sup>1</sup> were used in the order named, but in the first and second grades the latter form was omitted because of the fatiguing nature of the task.

The list of accepted responses given by Whipple was followed exactly and the results of the two tests were averaged for the final score for all grades above the second.

<sup>2</sup>Whipple, 1915, Manual of Mental and Physical Tests, pp. 275 - 276.

### Anagrams

The anagrams test consisted of twenty-one 3 to 7 letter words taken from the lower end of Ayres' spelling scale. A definite scheme of letter displacement was followed and the words were arranged in order of difficulty estimated from their relative length and number of letters displaced. This trial list was given to the first, second, sixth and eighth grade pupils of another school and the words were then arranged in increasing order of difficulty as shown by a frequency table derived from over one hundred cases. In this test, spelling was considered an essential element of the correct response, but spelling ability was minimized as a factor by the use of the easiest words from the Ayres list. To offset the puzzle element as much as possible, the subjects were instructed, in case they came to a word they could not get, to go on to the next and to come back to the hard word later.

### Cancellation

A Cancellation test was included in the list for purposes of comparison. A sheet of pied type, devised for the purpose, was used. Three letters were cancelled.

### Comprehension

For the comprehension test, three of the best individual tests from the Kuhlmann 1917 Revision were combined and adapted for group work. Two test sheets were used, the easier for all grades, the harder for the 4th to the 8th grades inclusive. Where results from both forms were used, the two sheets were added to

give the total score.

### Language Completion

The Trabue test was selected as the most thoroughly standardized language completion test for grade school subjects. The procedure and instructions were modified to suit the purpose of the present work. A trial test was given to all grades. Scale B also was given to all grades and scale L to grades 2A to 8B inclusive. The scores of an individual for the two tests were added to secure his total score. Responses listed by Trabue were the only ones allowed credit.

## Chapter III

## Criticism of Tests

The present tests are open to all the criticism that applies to group tests in general. The most serious, perhaps, is that the mood of the subject and hence the attention and amount of effort by him cannot be controlled by the experimenter to the same extent as when testing him individually. It is this fact, presumably, that accounts for the occasional instances where subjects "go to pieces" and obtain scores that are entirely out of harmony with their school records, and teachers' estimates and with the results of individual testing. It is this fact also that makes it impossible to rely upon group test results for a final diagnosis, and makes it necessary to use them merely to indicate where individual testing can be done to good advantage in sorting out supernormal and subnormal subjects from the mass of individuals of average ability.

That some cheating occurred in spite of the efforts to guard against it, is an undoubted fact. However, it is highly questionable that the average child who will cheat is capable of cheating intelligently.

The amount of coaching that may have been done and its possible effect upon either individual or total scores is impossible to determine.

A further criticism of group tests is that a test for a given ability does not test that ability unless the subject understands the task. There is no way by which the experimenter can make sure that every subject knows what he is to do.

Of the type of test used in the present work it can be said that the statement that the same tests are given to all ages is the purest sophistry, for though the situation may be the same, the response undoubtedly involves a very different kind of ability on the part of a 6 year old from that tested in a 13 year old subject. The digits test, for instance, probably affords a measure of immediate memory for the older subjects, but for a six year old with but little experience in writing digits it involves a large number of abilities of which ingenuity is, perhaps, not the least.

The answer to both of these latter criticisms is the same: the object of the present work is not to test any one ability or group of abilities, but to measure general intelligence. The fact that attention and comprehension have a more prominent role than in individual testing certainly does not detract from the value of group test results nor does the involving of different abilities at different age levels militate against the use of the tests for all ages as tests of general ability.

In addition to the defects common to all group tests, the tests used in this work have the specific defects of overstressing speed and language ability, and of being too dependent upon school attainment. In four of the nine tests the time limit is 100 seconds or less. In these, the slow but capable child scarcely gets started before it is time to stop. Although the result is not a fair index of his intelligence, it is a good illustration of what frequently happens to him in group work. The group test serves to mark him for further testing and individual help.

The stressing of language ability, though probably not



a defect in the tests for children from homes where English is the only language spoken, is a serious handicap to their use in schools where there is a large foreign element.

Whether the dependence upon school attainment is a real defect is a mooted question. On the one hand, it would seem to interfere with the measurement of native endowment. On the other, the amount of reading, writing, and spelling ability possessed by the subject would appear to bear some relation to his general ability. The tests, however, are found to give but slight indication of the real ability of superior 6 year olds owing to their lack of school training. They score better than average 6 year olds, but cannot get the score of the mental age to which they are entitled.

It is evident that the norms given are based upon too few cases to make them wholly reliable. With the addition of more cases the steps between successive ages would presumably tend to become equalized. For most of the tests there is at least one point at which the increase from one age to the next is too small to be significant. A larger number of cases would show whether this was due to the nature of the test or to the chance variation of the groups tested.

## Chapter IV

## Discussion of Data

The mean, as the most probable true average, was taken as the norm. This was found to increase, even though but slightly, from one age to the next for all tests excepting cancellation for which the 9 and 10 year scores are identical.

The language completion, anagrams, comprehension and cancellation tests were given after the advanced 13 year olds had entered high school. The rest of the 13 year group seems to have done well in cancellation, but to have done but little better than average 12 year olds on the other three tests. For those tests, therefore, it is probable that the present 13 year norms are too low. Judging from the table showing the correlation of chronological age with school grade, it would seem that the 10 year group had fewer bright subjects than had the other age groups. If this were true, the 10 year norms for the various tests would probably be raised by the addition of a larger number of cases.

Owing to differences in the form of the test used in some cases and in the procedure in others, most of the results secured are not directly comparable with norms already established.

For digits, the norms seem to run somewhat low.

## Digits

Age	3	4	5	6	7	8	9	10	11	12	13
Kuhlmann <sup>1</sup>		3			5						
Terman <sup>2</sup>	3	4			5			6			
Smedley <sup>3</sup>					5	5	5	6	6	6	6
Norm				3	4.37	4.86	5.26	5.39	5.74	5.88	5.98

<sup>1</sup>Whipple. 1915 Manual of mental and physical tests

<sup>2</sup>Terman, 1916, The measurement of intelligence.

<sup>3</sup>Kuhlmann, 1917, The measurement of mental development.

Table I

## Norms

## Easy Opposites

Age	6	7	8	9	10	11	12	13
No. of Cases		86	82	84	98	95	87	71
Av.	0	1.1	2.6	4.4	5.2	6.5	7.4	8.6
Stand. Dev.		1.2	1.5	2.5	2.1	1.9	2.1	2.6

## Hard Opposites

Age	6	7	8	9	10	11	12	13
No. of Cases		86	82	84	97	95	87	71
Av.	0	.5	1.3	2.8	3.7	6.3	7.1	9.1
Stand. Dev.		.98	1.5	3.0	3.1	4.4	4.1	4.5

## Substitution

Age	6	7	8	9	10	11	12	13
No. of Cases	68	88	82	84	98	94	87	71
Av.	10.1	20.5	30.2	40.3	44.2	50.6	52.3	59.1
Stand. Dev.	6.2	11.1	11.5	12.0	12.5	11.2	10.7	13.1

## Word Building

Age	6	7	8	9	10	11	12	13
No. of Cases	67	86	82	84	98	95	87	71
Av.	.4	2.3	4.3	6.3	7.4	9.8	10.3	12.0
Stand. Dev.	.9	2.5	2.8	3.4	3.2	3.9	3.7	3.0

Table I (cont.)  
Digits

Age	6	7	8	9	10	11	12	13
No. of Cases	66	87	82	83	98	95	87	71
Av.	3.0	4.37	4.86	5.26	5.39	5.74	5.88	5.98
Stand. Dev.	1.5	.75	.69	.81	.81	.77	.78	.96

## Language Completion

Age	6	7	8	9	10	11	12	13
No. of Cases	66	85	79	78	92	90	84	60
Av.	1.7	5	7.5	10.3	11.4	14.2	15.0	15.1
Stand. Dev.	2.2	3.0	3.1	4.4	3.6	4.7	3.8	4.4

## Anagrams

Age	6	7	8	9	10	11	12	13
No. of Cases	70	90	82	83	98	95	87	58
Av.	1.8	4.5	7.4	9.1	9.2	10.2	11.1	11.7
Stand. Dev.	2.2	3.5	3.1	3.1	3.2	2.8	2.8	3.4

## Cancellation

Age	6	7	8	9	10	11	12	13
No. of Cases	62	71	73	71	86	91	84	56
Av.	12.1	16.0	17.2	20.6	20.6	23.0	23.7	25.6
Stand. Dev.	5.2	4.7	3.9	5.0	4.9	5.1	5.7	5.6

## Comprehension

Age	6	7	8	9	10	11	12	13
No. of Cases	57	76	72	72	84	85	79	53
Av.	2.5	4	5.8	9.5	13.1	16.1	16.8	17.2
Stand. Dev.	2.5	2.0	3.9	6.2	5.3	3.9	3.3	2.9

For the lower groups, the effect of writing the response doubtless acted as a distraction. Averaging the results for the two series instead of taking the best of three trials, the method in individual tests, must also have tended to lower the norms for all ages. It seems reasonable to assume that the difference between the results and those of other investigators is due to the differences in procedure and method of scoring.

Although the results of the word building test, in which the final score was secured by averaging the results of the two trials, are not wholly comparable with those of Anderson and Pyle<sup>1</sup> and of Pintner<sup>2</sup> for the a e i r l p test alone, such a comparison is interesting.

#### Word Building

Age	6	7	8	9	10	11	12	13
Anderson Boys' Av.			5.5	7.3	8.3	10.6	11.5	12.6
and Pyle Girls' Av.			6.5	7.7	10.2	11.5	13.3	14.7
Pintner Med.	.0	1	3	4	5	6	8	9
Norm	.4	2.3	4.3	6.3	7.4	9.8	10.3	12.

To deal with the data effectively by the "more or less" method, a means had to be found of reducing the norms for the different tests to comparable units which should be distributed according to the discriminative value of the various tests, and for combining these units to form a scale.

<sup>1</sup>Whipple, 1915, Manual of Mental and Physical Tests.

<sup>2</sup>Pintner, p. 277, 1918, Mental Survey, p. 20.

## Chapter V

## Treatment of Data

The discriminative value of a test consists in the capacity of the test to classify individuals according to the ability tested. To find the discriminating value of each test for each age group, the difference between the average score for that age and that of the next lower age group was divided by the average of the standard deviations of the two groups.  $(Dif / (S1 + S2) / 2)$

This quotient represents the portion of the distribution area included between the average of one year and that of the next year either below or above it. Translated into terms of percentage, it is one half the percent of the group which scores neither so low as the average of the next lower age group, nor so high as the average of the next higher age group.

A large quotient indicates that a given test has served to mark off clearly a large proportion of the individuals of a given age group from those of older and younger groups: i. e. the test has shown a high degree of discriminative value.

A small quotient indicates that many of the individuals of a given age have made scores no better than the average of the next lower age group, and that many others have made scores as good as the average of the next higher age group. There is much overlapping of the successive groups and such tests are said to possess but little discriminative value.

Although differing in size for different tests and for different age groups of the same test, standard deviations are assumed always to represent equal psychological magnitudes. Quotients secured by the above formula being functions of standard

Table II

## Easy Opposites

	7-8	8-9	9-10	10-11	11-12	12-13
% Pass	13.9	13.4	32.1	22.4	31.5	37.6
Dif/S1*	10.6	11.6	37.5	27.1	34.5	27.5
% Fail	20.7	21.4	38.7	27.3	39.7	39.4
Dif/S2	15.9	23.6	35.2	25.2	35.2	30.9
% Av.	17.3	17.4	35.4	24.8	35.6	38.5
Dif/(S1+S2)/2 in %	13.4	18.5	36.7	25.8	34.5	29.2
% D. V.	.94	.94	.37	.68	.37	.29
Dif/(S1+S2)/2	1.1	.9	.34	.65	.40	.55

% Pass - % of lower group reaching Av. of higher group, actually found.

Dif/S1 - Same based on Dif/S1.

% Fail - % of higher group failing Av. of lower group, actually found.

Dif/S2 - Same based on Dif/S2.

% Av. - Av. of the above actually found percents.

Dif/(S1+S2)/2 - Same based on Dif/(S1+S2)/2  
in % in %

% D. V. - Discriminative value of Test based on Av. of found percents.

Dif/(S1+S2)/2 - Same based on Dif/(S1+S2)/2

\*S1 - Standard deviation of lower of groups compared  
S2 - " " " higher " " "

Table III

## Hard Opposites

	7-8	8-9	9-10	10-11	11-12	12-13
% Pass	12.7	18.2	23.7	16.3	40.	25.
Dif/S1	20.9	15.9	38.2	20.3	42.9	31.6
% Fail	36.5	40.4	42.2	30.5	47.7	39.4
Dif/S2	29.8	30.9	38.6	27.8	42.5	33.9
% Av.	24.6	29.3	32.9	23.4	43.8	32.2
Dif/(S1+S2)/2 in %	26.4	26.5	38.2	24.5	42.9	32.3
% D. V.	.69	.55	.44	.73	.16	.46
Dif/(S1+S2)/2	.63	.66	.295	.69	.18	.46

Table IV

## Word Building

	6-7	7-8	8-9	9-10	10-11	11-12	12-13
% Pass	4.4	23.2	23.1	38.0	21.4	44.2	35.2
Dif/S1	1.7	21.2	23.9	37.5	22.7	44.9	32.7
% Fail	16.2	25.3	27.8	39.7	31.5	48.8	35.2
Dif/S2	22.4	23.9	28.1	36.7	27.1	44.9	32.3
% Ave	10.3	24.3	25.4	38.8	26.4	46.5	35.2
% Av.							
Dif/(S1+S2)/2 in %	13.6	22.7	26.2	37.1	25.2	44.9	30.9
% D. V.	1.3	.70	.66	.28	.63	.09	.38
Dif/(S1+S2)/2	1.1	.75	.64	.33	.67	.13	.50



Table V

## Substitution

	6-7	7-8	8-9	9-10	10-11	11-12	12-13
% Pass	5.8	15.9	14.6	36.9	27.5	41.4	23.8
Dif/S1	4.7	19.3	19.3	39.5	30.5	44.1	26.5
% Fail	15.9	15.8	19.0	42.8	27.6	50.0	26.7
Dif/S2	17.4	20.1	20.1	37.9	28.5	43.7	30.2
% Av.	10.9	15.85	16.8	29.8	27.55	45.7	25.2
Dif/(S1+S2)/2 in %	11.5	19.8	19.8	37.9	29.5	44.1	28.5
% D. V.	1.2	1.0	.96	.26	.60	.11	.67
Dif/(S1+S2)/2	1.2	.85	.85	.31	.54	.15	.57

Table VI

## Digits

	6-7	7-8	8-9	9-10	10-11	11-12	12-13
% Pass	16.6	23.2	23.1	38.0	21.4	44.2	35.2
Dif/S1	18.1	21.2	23.9	37.5	22.7	44.9	32.7
% Fail	5.6	25.3	27.8	39.7	31.5	48.8	35.2
Dif/S2	3.4	23.9	28.1	36.7	27.1	44.9	32.3
% Av.	11.2	24.25	25.4	38.8	26.4	46.5	35.2
Dif/(S1+S2)/2 in %	11.5	22.7	26.2	37.1	25.2	44.9	30.9
% D. V.	1.2	.70	.66	.28	.63	.09	.38
Dif/(S1+S2)/2	1.2	.75	.64	.33	.67	.13	.50

Table VII

## Language Completion

	6-7	7-8	8-9	9-10	10-11	11-12	12-13
% Pass	15.1	27.0	15.1	42.3	17.3	43.3	41.1
Dif/S1	6.7	20.3	17.6	40.1	22.1	43.2	48.8
% Fail	12.4	18.9	21.7	38.0	30.0	49.3	50.0
Dif/S2	13.6	21.2	26.4	38.2	27.8	41.7	49.2
% Av.	13.8	22.9	18.4	40.2	23.7	46.3	45.5
Dif/(S1+S2)/2 in %	9.7	21.2	21.2	38.2	24.2	42.1	49.2
% D. V.	1.0	.74	.9	.25	.71	.10	.11
Dif/(S1+S2)/2	1.3	.8	.8	.3	.7	.2	.02

Table VIII

## Anagrams.

	6-7	7-8	8-9	9-10	10-11	11-12	12-13
% Pass	15.7	24.4	24.3	43.3	32.6	30.5	43.6
Dif/S1	11.1	20.7	30.5	48.8	37.9	37.5	41.7
% Fail	26.7	13.4	24.1	52.0	46.5	45.9	46.5
Dif/S2	22.1	17.7	30.5	48.8	36.4	37.5	43.3
% Av.	21.2	18.9	24.2	47.7	39.6	38.2	45.1
Dif/(S1+S2)/2 in %	18.4	19.3	30.5	48.8	37.1	37.5	42.5
% D. V.	.9	.9	.7	.06	.3	.3	.12
Dif/(S1+S2)/2	.9	.9	.6	.03	.3	.3	.19

Table IX  
Cancellation.

	6-7	7-8	8-9	9-10	10-11	11-12	12-13
% Pass	30.6	42.2	21.9	47.8	31.4	47.5	34.5
Dif/S1	22.7	39.7	19.2	0	31.2	44.4	37.0
% Fail	11.3	36.9	25.3	56.9	25.2	51.2	44.6
Dif/S2	20.3	37.8	24.8	0	31.9	52.2	36.7
% Av.	21.0	39.6	23.6	52.4	28.3	49.4	39.6
Dif/(S1+S2)/2 in %	21.2	38.2	22.4	0	31.6	44.8	36.7
% D. V.	.8	.3	.7	-.1	.6	.02	.3
Dif/(S1+S2)/2	.8	.3	.8	0	.5	.1	.3

Table X  
Comprehension

	'6 - 7	'7 - 8	'8 - 9	'9 - 10	'10 - 11	'11 - 12	'12 - 13
Percent of lower group reaching average of higher group, actually found	19.3	19.7	9.7	31.9	28.5	55.7	46.8
<u>Difference</u>							
Same based on S 1	27.4	18.4	17.4	28.1	28.4	42.9	45.2
Percent of higher group failing average of lower group	25.0	50	40.2	28.6	22.3	41.7	38.9
<u>Difference</u>							
Same based on S 2	22.7	32.3	27.4	24.8	22.1	41.7	44.4
Average of above actually found percents	22.2	34.9	25.0	30.3	25.4	48.7	40.4
<u>Difference</u>							
Same based on $\frac{S 1 + S 2}{2}$	24.2	27.4	23.3	26.8	25.8	42.5	44.8
Discriminative value of test based on average of found percents	.7	.4	.7	.5	.7	.03	.1
<u>Difference</u>							
Same based on $\frac{S 1 + S 2}{2}$	.7	.6	.7	.6	.7	.2	.1

deviation may therefore be treated as comparable units.

Under these assumptions these quotients represent an absolute increase in ability from one age to the next for their respective tests. If the tests can be regarded as tests of intelligence, then these quotients are absolute increases of intelligence between the years concerned, and a scale built up from them can properly be regarded as a scale of absolute increases in intelligence beyond the starting point  $x$ .

This starting point  $x$ , or arbitrary zero has been placed for all tests at the six year average. The quotient representing the increase between the 6 year ability and 7 year ability in a given test is added to the  $x$  and is equivalent to the 7 year score for that test. Similarly, the quotient representing the increase from 7 year ability to 8 year ability for the same test indicates the number of units to be added to those corresponding to the 7 year score in order to determine the equivalent of the 8 year score. By continuing this method of procedure through all age groups of all the tests a scale is constructed the unit of which is a quotient of one or standard deviation.

For those tests showing a high degree of discriminative capacity for all ages more units are thus gradually built up than for those showing but a small absolute increase from age to age.

By considering these units as points and combining the points of the various tests for each age group a total score can be secured corresponding to each age from 6 to 13. But at this point a change in definition must be made. The subjects we have been regarding as 6 year old, i. e., from 6:0 to 6:11 inclusive have an average age of 6:6. Hence, the 6 year score must be placed

Table XI

## Points Corresponding to Norms

Years		6	7	8	9	10	11	12	13
Opposites I	Av.	0	1.1	2.6	4.4	5.2	6.5	7.4	8.6
	Points X		.9	2.0	2.9	3.2	3.9	4.3	4.9
Opposites II	Av.	0	.5	1.3	2.8	3.7	6.3	7.1	9.1
	Points X		.5	1.1	1.8	2.1	2.8	3.0	3.4
Substitution	Av.	10.1	20.5	30.2	40.3	44.2	50.6	52.3	59.1
	Points X		1.2	2.1	3.0	3.3	3.8	4.0	4.6
Word Building	Av.	.4	2.3	4.3	6.3	7.4	9.8	10.3	12.0
	Points X		1.1	1.9	2.5	2.8	3.5	3.6	4.1
Digits	Av.	3	4.37	4.86	5.26	5.39	5.74	5.88	6.
	Points X		1.2	1.9	2.4	2.6	3.0	3.2	3.3
Trabue	Av.	1.7	5	7.5	10.3	11.4	14.2	15.0	15.1
	Points X		1.3	2.1	2.9	3.2	3.9	4.6	4.1
Anagrams	Av.	1.8	4.5	7.4	9.1	9.2	10.2	11.1	11.7
	Points X		.9	1.8	2.3	2.3	2.7	3.0	3.2
Cancellation II	Av.	12.1	16.0	17.2	20.6	20.6	23.0	23.7	25.6
	Points X		.8	1.1	1.9	1.9	2.4	2.5	2.8
Comprehension	Av.	2.5	4	5.8	9.5	13.1	16.1	16.8	17.2
	Points X		.7	1.3	2.0	2.6	3.3	3.5	3.6
Total Points		X	8.6	15.3	21.7	24.0	29.3	31.2	34.0

Table XII

Points for Successive Ages According to Combined Curve

	0 mo.	1 mo.	2 mo.	3 mo.	4 mo.	5 mo.	6 mo.	7 mo.	8 mo.	9 mo.	10 mo.	11 mo.
5 yrs.	X											
6 yrs.							X	.7	1.4	2.2	2.9	3.5
7 yrs.	4.3	5.0	5.8	6.5	7.2	7.9	8.6	9.2	9.7	10.3	10.8	11.4
8 yrs.	12.0	12.5	13.1	13.6	14.2	14.8	15.3	15.8	16.4	16.9	17.4	18.0
9 yrs.	18.5	19.0	19.5	20.1	20.6	21.1	21.7	21.9	22.1	22.3	22.5	22.7
10 yrs.	22.8	23.0	23.2	23.4	23.6	23.8	24.0	24.4	24.9	25.3	25.8	26.2
11 yrs.	26.6	27.1	27.5	28.0	28.4	28.8	29.3	29.5	29.6	29.8	29.9	30.1
12 yrs.	30.3	30.4	30.6	30.7	30.9	31.1	31.2	31.4	31.7	31.9	32.1	32.4
13 yrs.	32.6	32.8	33.0	33.3	33.5	33.7						

Fig. 1. Easy Opposites

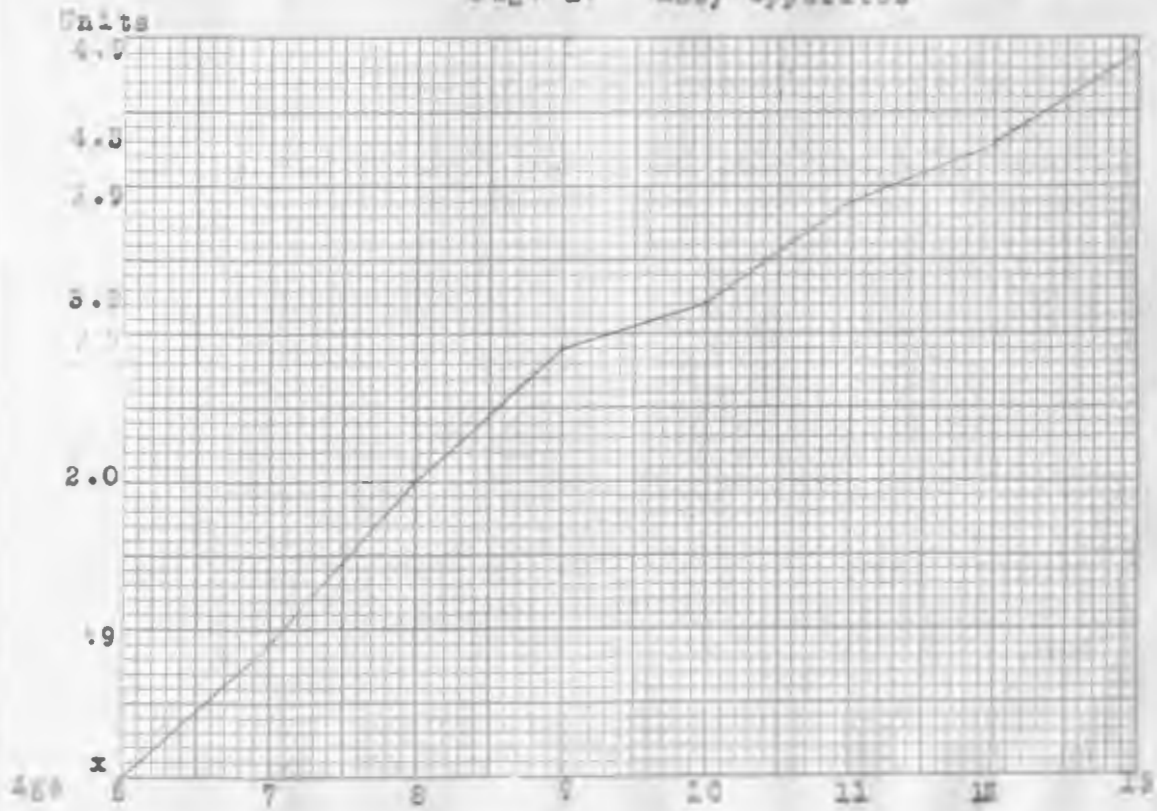
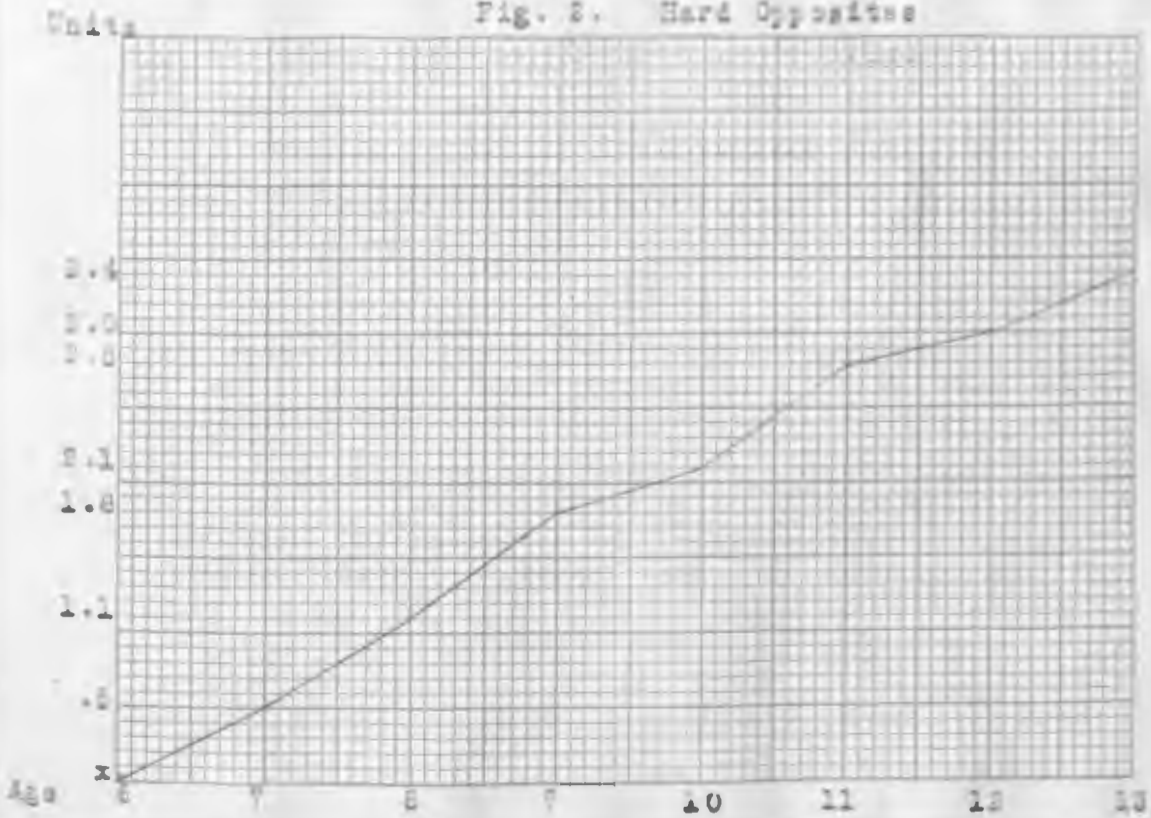


Fig. 2. Hard Opposites





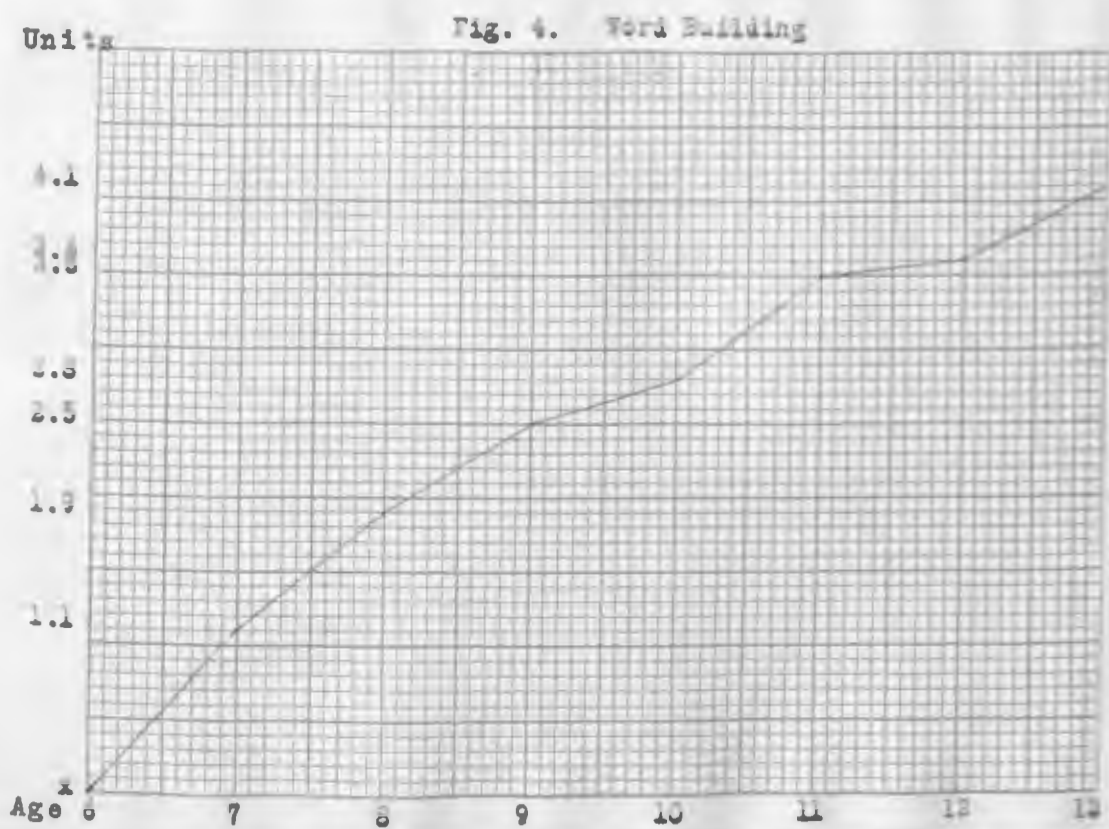
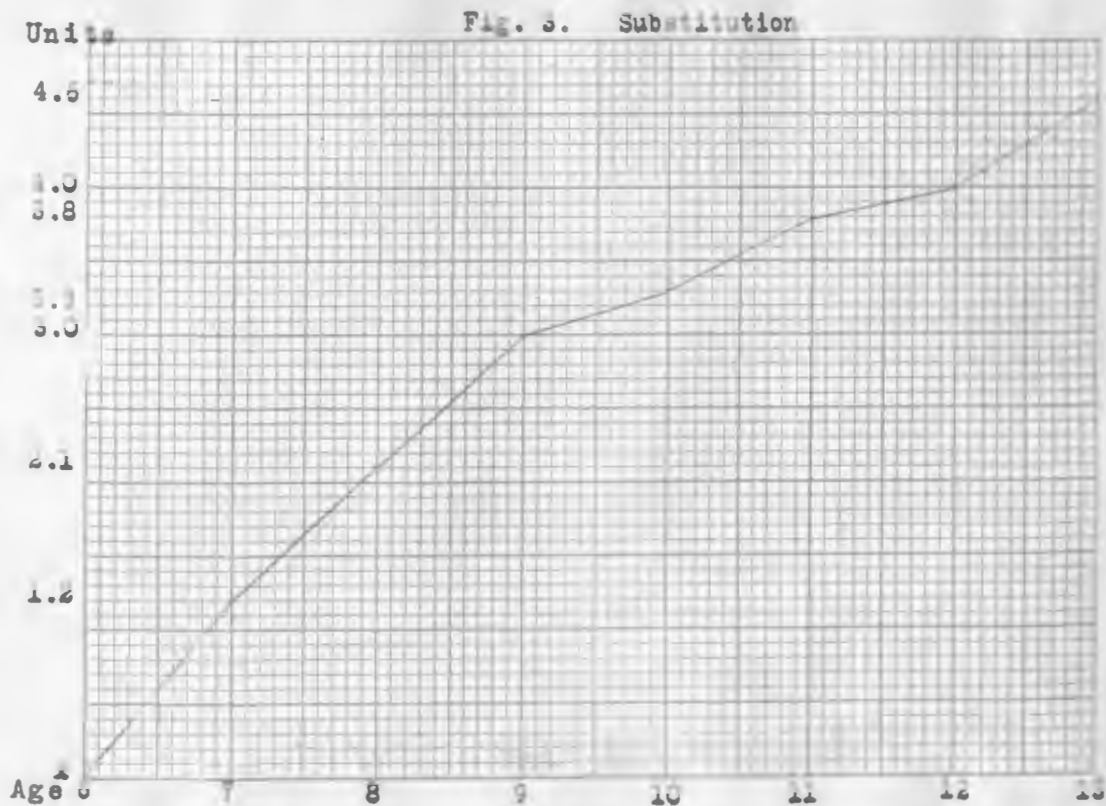


Fig. 5. Digits

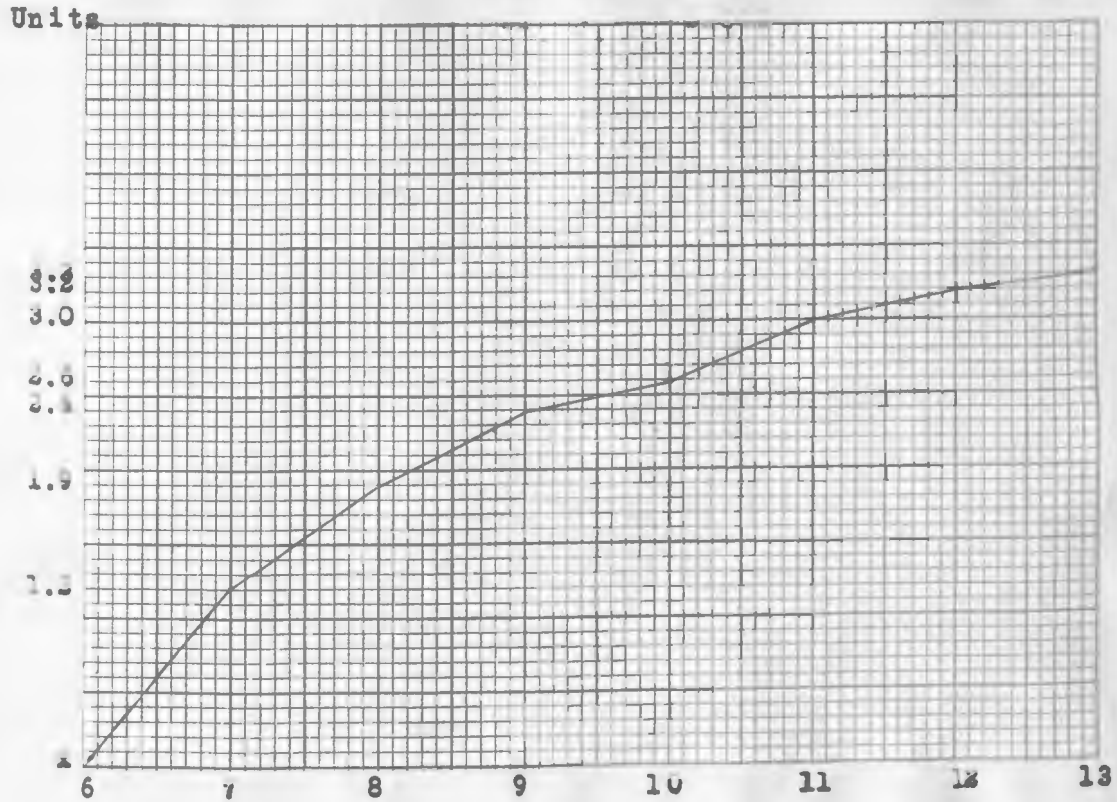
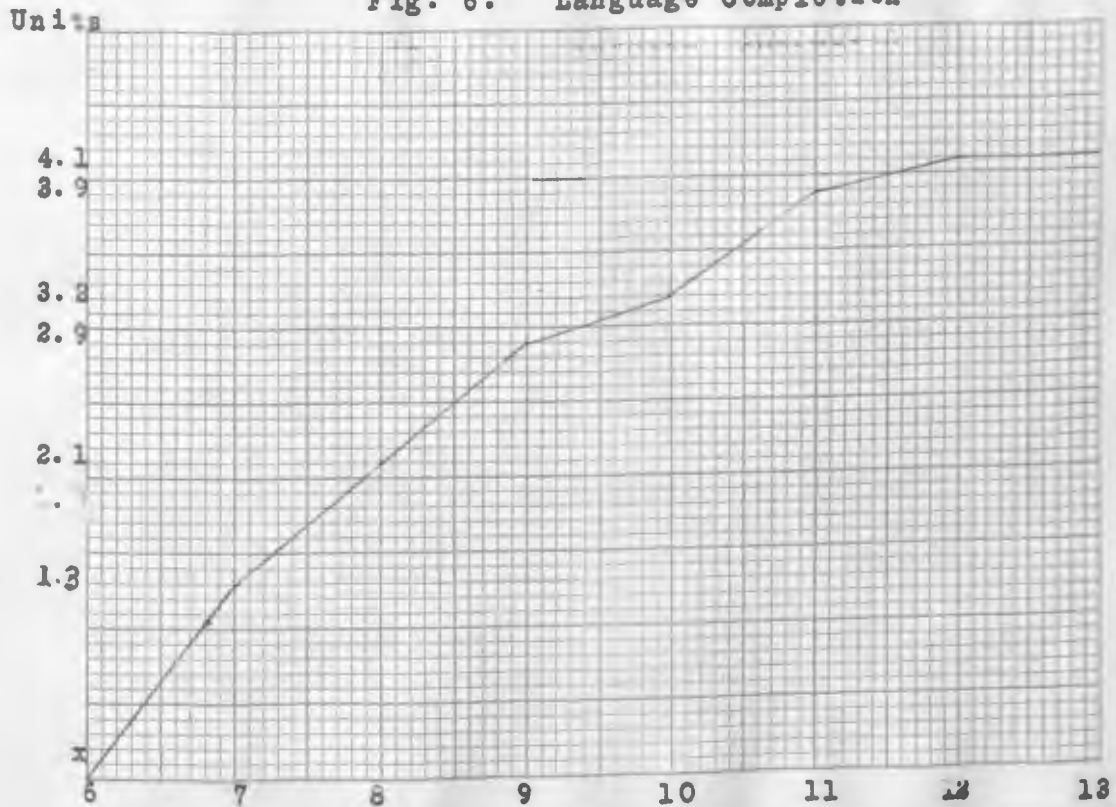


Fig. 6. Language Completion



Rate of Development from Age to Age as Shown by Separate Tests.

Fig. 7. Anagrams

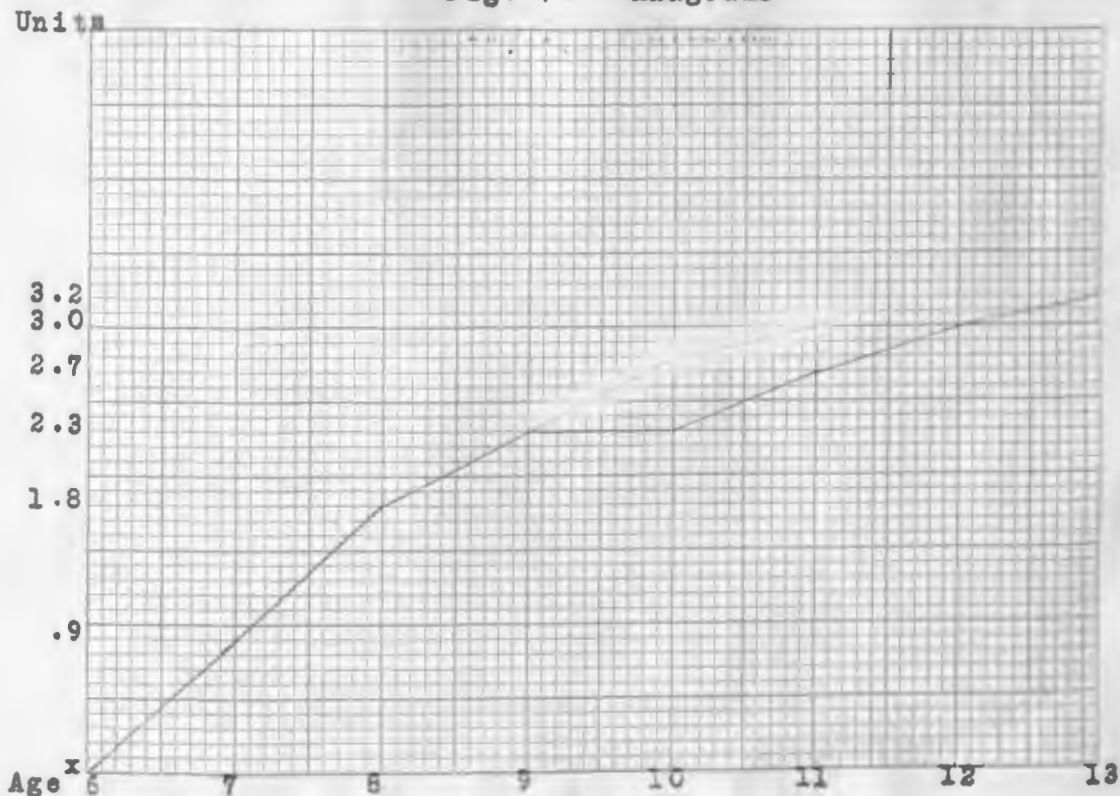


Fig. 8. Cancellation

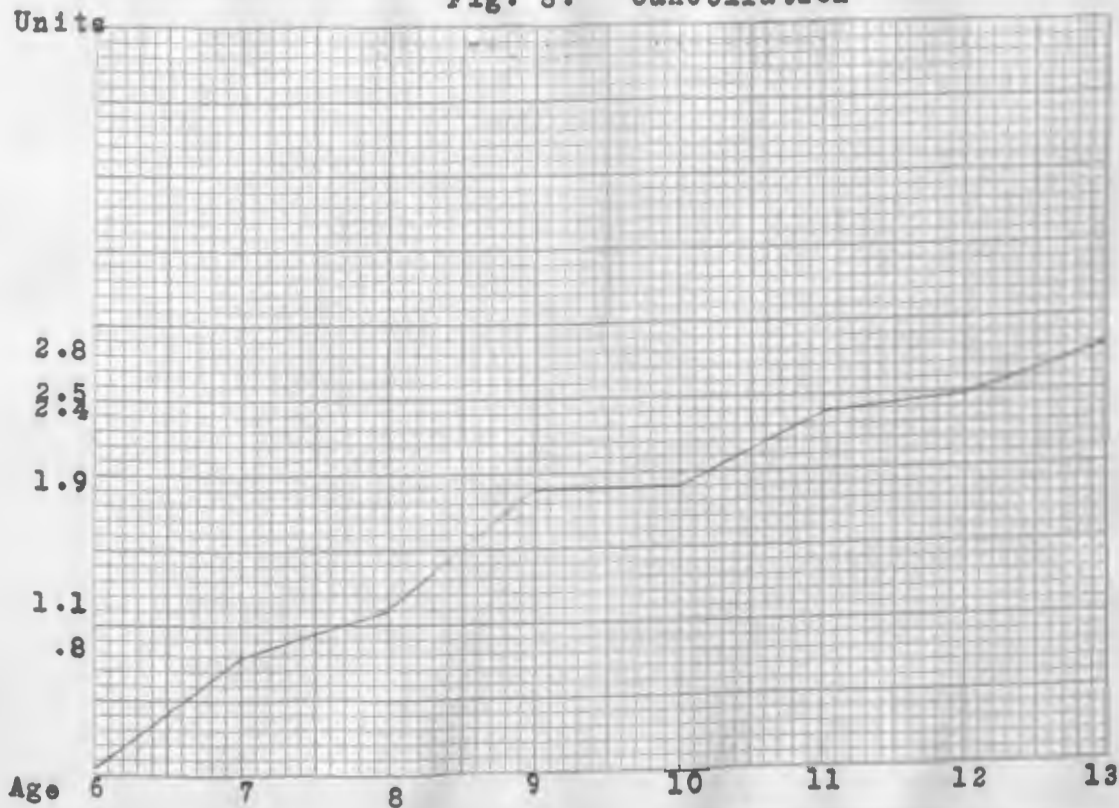
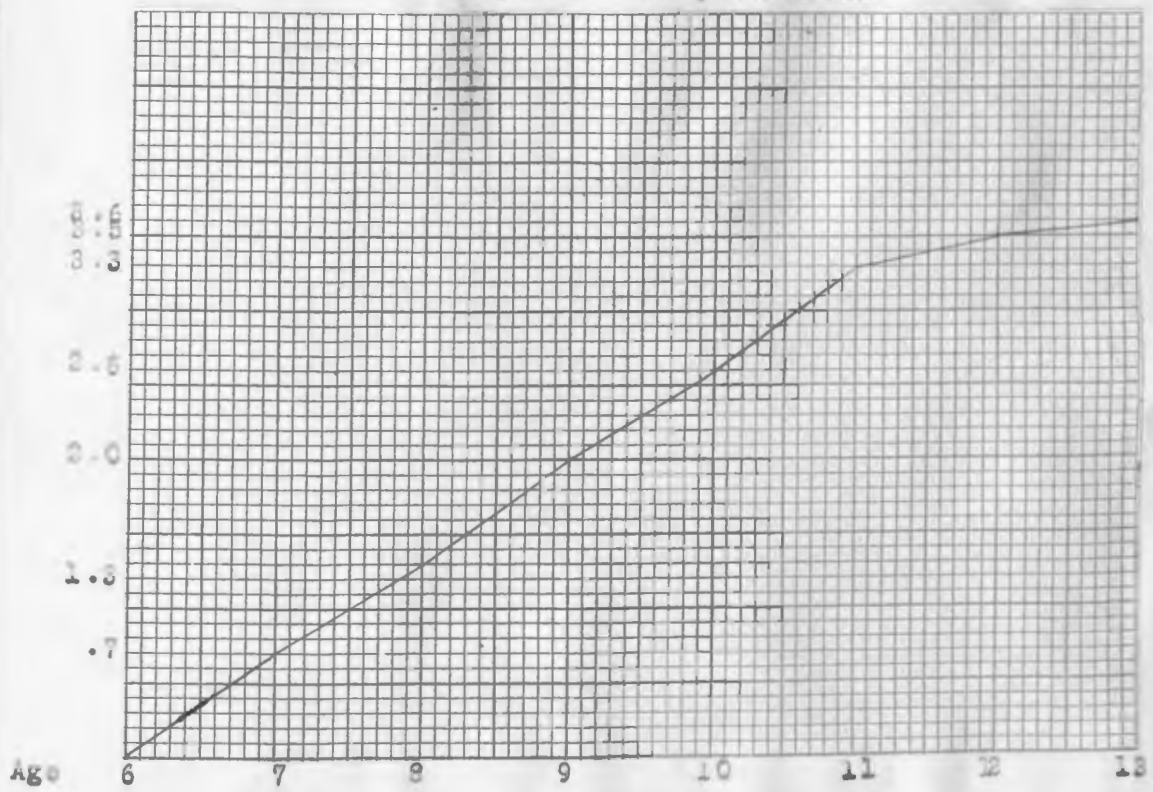
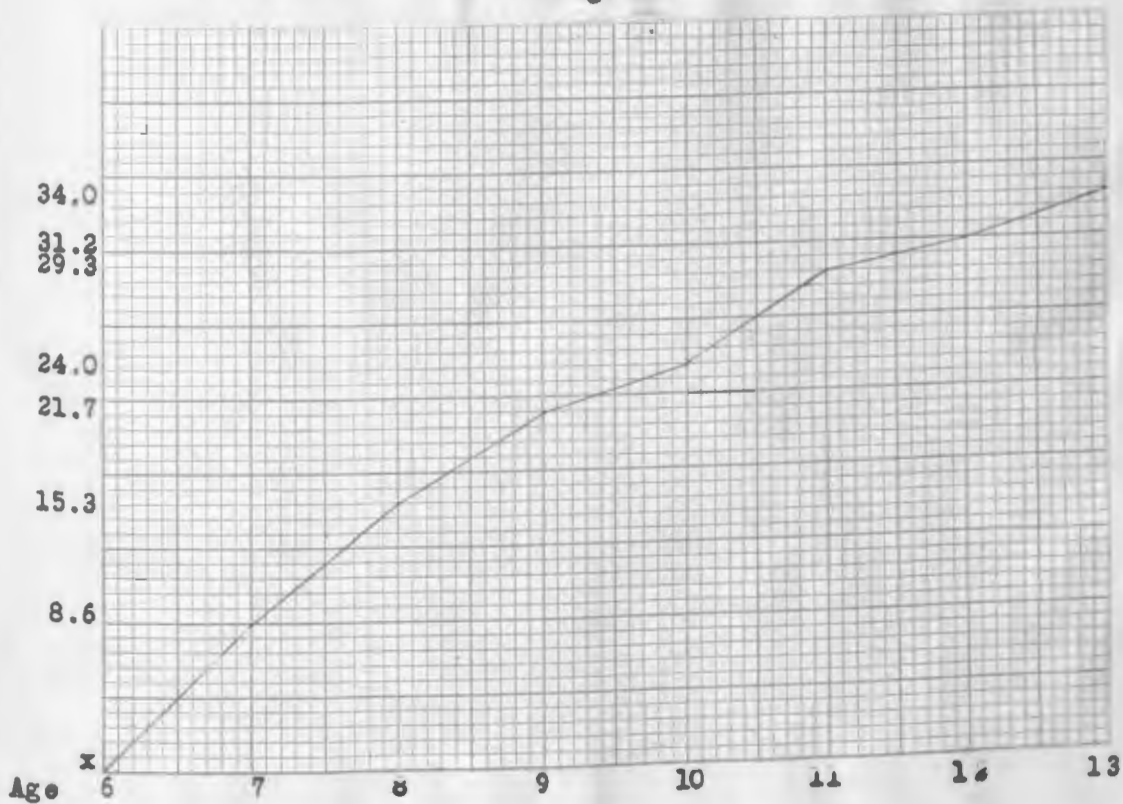


Fig. 9. Comprehension



Curve Showing the Rate of Development of Absolute Intelligence from Age to Age, as Based Upon the Combined Results of the Separate Tests.

Fig. 10.



Drawn on scale of 1 to 9.

at the 6.6 point on the combined curve if the norm is to correspond to the age to which it really belongs. Because of the different rate of increase of absolute intelligence units or points from age to age in the different tests there takes place in the combined score an automatic weighting in favor of the more valuable tests.

The subject who makes a 13 year score in the easy opposites will earn 4.9 points (Fig. 1 ) while a subject who makes a 13 year score in the cancellation will earn only 2.8 points, (Fig. 7) not because the former is regarded as a good test and the latter is held in small esteem but because the one, by exhibiting a higher degree of discriminative capacity for each age group, has added for each age a larger number of units or points than the other.

In the present work, the discriminative value of the test as derived from  $\text{Dif.} / ((S1+S2)/2)$  has been used in preference to that based upon the found percents, under the assumption that it is more reliable owing to the larger number of cases involved. In the majority of instances, however, the results derived by the two methods are so nearly identical as to make it feasible to use whichever is more convenient. In further work, therefore, while norms are shifting with the addition of new cases and until permanent norms are established, it will probably be advisable to avoid having to find the standard deviation every time a change in the norm is made by using the discriminative value based upon percents.

When the number of points, based upon the discriminative value of a test as derived from the formula  $\text{Dif.} / ((S1+S2)/2)$ , had

Table XIII

Easy Opposites	Substitution	Word Building	Digits	Cancellation
Score Points	Score Points	Score Points	Score Points	Score Points
0 X	0 -1.6	0 -.4	0 -2.0	0 -2.4
.5 .4	1 -1.5	.5 .1	.25 -1.8	1 -2.2
<u>1.0 .8</u>	2 -1.3	1.0 .4	.5 -1.7	2 -2.0
1.5 1.2	3 -1.1	1.5 .7	.75 -1.5	3 -1.8
2.0 1.6	4 -1.0	2.0 .9	1.00 -1.3	4 -1.6
2.5 1.9	5 - .8	2.5 1.2	1.25 -1.2	5 -1.4
3.0 2.2	6 - .7	3.0 1.4	1.50 -1.0	6 -1.2
3.5 2.5	7 - .5	3.5 1.5	1.75 - .8	7 -1.0
4.0 2.7	8 - .3	4.0 1.7	2.0 - .7	8 - .8
4.5 2.9	9 - .2	4.5 2.0	2.25 - .5	9 - .6
5.0 3.2	<u>10 X</u>	5.0 2.1	2.50 - .3	10 - .4
5.5 3.4	11 .1	5.5 2.3	<u>2.75 - .2</u>	11 - .2
6.0 3.6	12 .2	6.0 2.4	3.0 X	<u>12 X</u>
6.5 3.9	13 .3	6.5 2.6	3.25 .2	13 .2
7.0 4.2	14 .4	7.0 2.7	3.50 .4	14 .4
7.5 4.3	15 .5	7.5 2.8	3.75 .6	15 .6
8.0 4.6	16 .7	8.0 3.0	4.00 .8	16 .8
<u>8.5 4.8</u>	17 .8	8.5 3.1	4.25 1.0	17 1.0
9.0 5.0	18 .9	9.0 3.2	4.50 1.4	18 1.3
9.5 5.2	19 1.0	9.5 3.4	4.75 1.7	19 1.5
10.0 5.4	20 1.1	10.0 3.6	5.00 2.1	20 1.7
10.5 5.6	21 1.2	10.5 3.7	5.25 2.4	21 2.0
11.0 5.8	22 1.3	11.0 3.8	5.50 2.7	22 2.2
11.5 6.0	23 1.4	11.5 3.9	5.75 3.0	23 2.4
12.2 6.2	24 1.5	12.0 4.1	6.00 3.3	24 2.6

Table XIII (cont.)

Easy Opposites		Substitution		Word Building		Digits		Cancellation	
Score	Points	Score	Points	Score	Points	Score	Points	Score	Points
12.5	6.3	25	1.6	12.5	4.2	6.25	3.6	25	2.7
13.0	6.4	26	1.7	13.0	4.3	6.5	3.8	26	2.9
13.5	6.6	27	1.8	13.5	4.5	6.75	4.1	27	3.1
14.0	6.8	28	1.9	14.0	4.6	7.0	4.3	28	3.2
14.5	7.0	29	2.0	14.5	4.8	7.25	4.6	29	3.4
15.0	7.2	30	2.1	15.0	5.0	7.5	4.8	30	3.6
15.5	7.4	31	2.2	15.5	5.2	7.75	5.1	31	3.8
16.0	7.6	32	2.3	16.0	5.3	8.0	5.3	32	4.0
16.5	7.8	33	2.3	16.5	5.5	8.25	5.6	33	4.1
17.0	8.0	34	2.4	17.0	5.6	8.5	5.8	34	4.3
17.5	8.2	35	2.5	17.5	5.8	8.75	6.1	35	4.5
18.0	8.4	36	2.6	18.0	6.0	9.0	6.3	36	4.7
18.5	8.6	37	2.7	18.5	6.2			37	4.8
19.0	8.8	38	2.8	19.0	6.3			38	5.0
19.5	9.0	39	2.8	19.5	6.5			39	5.2
20.0	9.2	40	2.9	20.0	6.5			40	5.4
		41	3.0	20.5	6.8			41	5.6
		42	3.1	21.0	7.0			42	5.7
		43	3.2	21.5	7.2			43	5.9
		44	3.3	22.0	7.3			44	6.1
		45	3.4	22.5	7.5			45	6.3
		46	3.5	23.0	7.6				
		47	3.6						
		48	3.7						
		49	3.7						
		50	3.8						



Table XIII (cont.)

Hard Opposites	Language Completion	Comprehension	Anagrams
Score Points	Score Points	Score Points	Score Points
<u>0</u> <u>X</u>	0      -.8	0      -.8	0      -.8
1      .9	<u>1</u> <u>-.3</u>	1      -.4	<u>1</u> <u>-.4</u>
2      1.5	2      .1	<u>2</u> <u>-.2</u>	2      .1
3      1.9	3      .5	3      .2	3      .4
4      2.1	4      .9	4      .7	4      .8
5      2.4	5      1.3	5      1.0	5      1.1
6      2.7	6      1.6	6      1.3	6      1.4
7      3.0	7      1.9	7      1.5	7      1.7
8      3.2	8      2.2	8      1.7	8      2.0
<u>9</u> <u>3.4</u>	9      2.5	9      1.9	9      2.3
10      3.6	10      2.8	10      2.1	10      2.6
11      3.8	11      3.1	11      2.3	<u>11</u> <u>3.0</u>
12      4.0	12      3.3	12      2.4	12      3.3
13      4.2	13      3.6	13      2.6	13      3.6
14      4.4	14      3.8	14      2.8	14      3.9
15      4.6	<u>15</u> <u>4.1</u>	15      3.0	15      4.2
16      4.8	16      4.3	16      3.2	16      4.5
17      5.0	17      4.5	<u>17</u> <u>3.6</u>	17      4.8
18      5.2	18      4.7	18      3.9	18      5.1
19      5.4	19      4.9	19      4.2	19      5.4
20      5.6	20      5.1	20      4.5	20      5.7
	21      5.3	21      4.8	21      6.0
	22      5.5	22      5.1	
	23      5.7	23      5.4	
	24      5.9		
	25      6.1		
	26      6.3		

been determined for every age group of every test, the next step was to find the number of points corresponding to the actual scores. This was done according to the same formula used in finding the number of points corresponding to the norms. The difference between a given score and the next lower norm was divided by the average of the standard deviations of the two age groups between the norms of which the score in question was found. The result was added to the number of points corresponding to the next lower norm. Points were assigned in this way for every possible score lying between the norms for 5:6 and 13:5.

How to deal with the scores beyond these limits was a serious problem. It was finally assumed that the 6 year standard deviation for each test would apply to all scores for that test below the 6 year average. There is an element of error in this assumption, as it is supposed that standard deviations tend to decrease with age. If this supposition is correct the result of using the 6 year standard deviation is to penalize failure to reach the 5:6 average less severely than if the actual standard deviations of the lower age groups were used. The difference between any score falling below the 6 year average, and the 6 year norm was divided by the 6 year standard deviation for the test and, as the arbitrary zero or X was placed at 6, the result was given a negative value. In this way the scale was extended down to zero for every test.

Scores falling above the norm for 13 years were dealt with in like fashion on the assumption that the standard deviation for 13 years would apply to all higher age groups. There is more reason for this assumption than for the preceding one.

as it is not likely that standard deviations tend to change much after the 13th year.

To determine the total score for each age, the points corresponding to the norms of each age group for the nine tests were added. Norms for months were interpolated between these year norms. At this point there is a slight reversion to the "all or none" method, which is made advisable by the nature of the material. It would be entirely possible to continue with the "more or less" method, and by further interpolations to express every obtained score in terms of years, months and days. Considering, however, the somewhat approximate nature of most results obtained by mental tests, whether given to groups or individually, this would seem to be an over-refinement of method.

Mental ages, therefore, were reckoned simply in years and months, and all fractions of months were discarded. Intelligence quotients were found in the usual manner, by dividing the mental age by the chronological age. Although a method was found which offers a tentative solution of the short scale problem for the separate tests, none was found for placing on the combined curve mental ages below and above those for which norms were secured. No attempt was made to work out such a method as the practical need of one in the grade schools where the tests are to be used is not great.

When norms are finally established for all ages to 13 years, 5 months the method suggested by Otis<sup>1</sup> might prove practicable. This is based upon the normal curve and assumes that

<sup>1</sup>Otis, S. Some Logical Aspects of the Binet Scale. Psychological Review, May 1916, Vol. 23, p. 176.

of the subjects of a given age group, the same proportion which fails to exceed the norm of the age group next below, will reach that of the age group next above.

By finding the proportion of 13 year olds that are mentally 12, the percent which will reach a mental age of 14 can be estimated and the score which is secured by this proportion of the 13 year group can be taken as the 14 year norm. In the same way an hypothetical 15 year norm may be determined by means of the 13 year olds who reach only the 11 year level, etc. The estimated norm for each successive year will be less reliable than that of the preceding year, however, as it will be based upon a fewer number of cases.



Table XV

## Correlation of Chronological Age with School Grade

	7:0 to 7:5	7:6 to 7:11	8:0 to 8:5	8:6 to 8:11	9:0 to 9:5	9:6 to 9:11	10:0 to 10:5	10:6 to 10:11	11:0 to 11:5	11:6 to 11:11	12:0 to 12:5	12:6 to 12:11	13:0 to 13:5	13:6 to 13:11
Grade Kg														
1B		1												
1A	19	7		1										
2B	17	3	1	1	1	1								
2A	6	22	9	6			2	1						
3B	2	14	7	12	6	3	1							
3A		15	17	8	4	4	1	1						
4B		3	11	11	3	8	3	1	1					
4A			1	8	15	10	9	2			1		1	
5B				4	7	9	9	3	2	4	1	2		
5A				4	10	13	26	15	9	5	1		1	
6B						1	3	12	15	5	3	3	3	
6A								4	13	9	8	2		
7B								6	10	17	13	7	1	
7A								2		6	14	11	4	
8B										1	4	5	6	
8A												3	13	
H.S.												5	8	

### Results

By the aid of Table XII point scores were secured for 490 subjects. These scores, by reference to Table XII, were converted into terms of mental age and thence, into intelligence quotients.

The correlation between these mental ages and school grade is shown in Table XIV. The correlation between school grade and chronological age is shown in Table XV. As would be expected, the correlation is found to be closer in the former, especially in the first four grades. In these lower grades during the past year a systematic effort has been made to place pupils according to their ability as indicated by results obtained with the Kuhlmann 1917 Revision.

Table XVI shows the correlation of mental ages, derived from these tests by the "more or less" method described above, with chronological age. It will be noted that for the upper age groups there is a tendency for the mode to drop a half year below the point where it would be expected to fall. It is possible that this is due to the system of weighting. As standard deviations tend to increase from one age to the next, and the differences between age groups tend to decrease, it follows that a subject is penalized for falling below a norm slightly more than he is rewarded for going an equal distance above it. The difference is not great, but because of it, if it had not been for other considerations, the average number of points earned by each age group would have been used as the basis of mental age classification instead of the combined curve.

**Table XVI**  
**Correlation of Mental Age with Chronological Age**

Mental Age	Chronological Age							Total Cases
	7:0 to 7:11	8:0 to 8:11	9:0 to 9:11	10:0 to 10:11	11:0 to 11:11	12:0 to 12:11	13:0 to 13:11	
Less than 6:6	6		1					7
6:6 - 6:11	16	1						17
7:0 - 7:11	30	19	6	2	1			58
8:0 - 8:11	17	34	21	11	1	1	2	87
9:0 - 9:11	3	12	18	23	11	3	3	73
10:0 -10:11		3	7	17	18	15	3	63
11:0 -11:11		1	4	18	14	17	5	59
12:0 -12:11			7	3	8	8	7	33
13:0 -13:5			1	1	9	5	9	25
More than 13:5			1	4	19	25	19	68
<b>Total Cases</b>	<b>72</b>	<b>70</b>	<b>66</b>	<b>79</b>	<b>81</b>	<b>74</b>	<b>48</b>	<b>490</b>



Table XVII

Correlation of Group Test I.Q.s with those of Kuhlmann 1917 Revision  
of the Binet-Simon Scale

I.Q. based on Group Tests	I.Q. based on Individual Test							Total	
	60-69	70-79	80-89	90-99	100-109	110-119	120-129		130-139
60-69	1		1						2
70-79	1	4							5
80-89		3	2	1					5
90-99			2	9	5	1			17
100-109				1	5	5	3		14
110-119						4	1		5
120-129						1	1	1	3
130-139									0
Total	0	2	10	12	11	11	6	1	52

The use of the combined curve, however, makes it possible to secure an approximate mental age for a subject from whose final score the results from one or more of the tests are lacking. The number of points corresponding to the norms of the tests in question are subtracted from the total number of points required for a complete score. In the present work, scores for the various tests were secured from nearly 700 subjects, but for only 490 of these were complete scores obtained. Because of this, the mental age norm determined by the combined curve have been adhered to.

The tendency of the results to run somewhat low is seen also in Table XVII, which shows the correlation between intelligence quotients derived from the group tests by the method described above, with those secured by the giving of individual tests to 52 subjects. In this case, however, it is perhaps only fair to explain that a large percentage of the subjects who were given individual tests were recommended for testing by the teachers because of their failure to respond in group work.

Attention should again be called to the fact that this is primarily a study in method, although the results secured are subject to the same limitations and open to the same criticisms as the results from group tests in general, this does not involve a criticism of the method of dealing with these results. The same method can be applied equally well to the testing of individuals and results can be secured which are free from the objections arising from the nature of work with groups.

The records of the 52 subjects represented in Table XVII were re-scored according to the "all or none" method. The

Table XVIII

Comparison of I. Q.s Derived by the "All or None" Method with Those Obtained by the "More or Less" from the Same Data.

S. G. - School Grade.                      A. or N. I.Q. - "All or None" I.Q.  
 M. A. - "All or None" Mental Age.      M. or L. I.Q. - "More or Less" I.Q.  
 Chr. A. - Chronological Age.            Ind. I.Q. - Individual I.Q.

Case	S. G.	M. A.	Chr. A.	A. or N. I.Q.	M. or L. I.Q.	Ind. I.Q.
1	4B	9-11	9-0	110	117	124
2	2B	6-11	9-7	72	75	75
3	3A	8-9	8-7	101	100	102
4	3A	7-9	8-6	91	91	102
5	4B	9-3	8-7	107	107	117
6	3B	8-3	8-1	89	94	107
7	3B	8-9	8-9	100	93	91
8	3B	8-1	8-8	93	100	103
9	3B	7-5	8-3	87	96	100
10	2A	8-6	8-0	106	108	111
11	2A	8-5	8-4	100	105	101
12	3A	9-7	8-10	108	109	105
13	2A	7-9	8-6	91	96	104
14	3A	8-1	8-1	100	105	119
15	3B	7-9	8-9	88	90	88
16	3B	7-9	8-3	94	96	93
17	2A	7-5	8-7	86	86	102
18	2A	8-10	7-1	96	101	128
19	2A	7-1	7-9	91	95	103
20	2A	7-9	7-9	100	106	111
21	3B	9-1	7-7	120	120	134
22	3B	8-3	7-7	108	113	115

Table XVIII (cont.)

Case	S. G.	M. A.	Chr. A.	A. or N. I.Q.	M. or L. I.Q.	Ind. I.Q.
23	3B	7- 9	7- 8	100	105	121
24	7A	11- 6	13- 7	84	88	99
25	6B	8- 7	13- 2	66	67	73
26	7B	11- 2	12- 5	90	91	82
27	7A	11- 6	12- 0	96	97	98
28	4A	88- 6	10-11	77	79	80
29	3A	8- 3	10- 3	80	79	80
30	4B	8- 3	10-10	76	79	86
31	4B	7- 9	10- 7	73	77	88
32	5B	9- 6	10- 4	91	91	92
33	5B	9- 1	10- 5	87	91	93
34	5B	10- 7	10-10	97	103	94
35	5B	9- 9	10- 6	93	93	91
36	5A	9- 4	10- 7	88	97	114
37	5A	9-10	10- 8	92	100	100
38	4B	9- 6	10- 1	94	93	98
39	3A	8- 7	10- 0	86	88	82
40	3A	8- 6	10- 2	83	83	87
41	2A	7- 5	10-11	67	68	85
42	5A	11- 2	9- 5	118	124	124
43	4B	8- 5	9- 0	93	96	97
44	3A	8- 2	9-11	82	88	92
45	4B	10- 5	9- 3	112	114	114
46	3B	9- 7	7- 8	123	121	111
47	3B	8- 7	7-11	108	108	116
48	3B	8- 2	7- 7	107	115	117

Table XVIII (cont.)

Case	S. G.	M. A.	Chr. A.	A. or N. I.Q.	M. or L. I.Q.	Ind I.Q.
49	2A	8- 5	7- 5	113	113	111
50	2B	7-10	7- 2	109	106	123
51	1A	6- 2	7-10	78	83	85
52	1A	6- 7	7- 5	88	93	98

norms were placed at the mid-point of the age groups from which they were derived, to avoid a constant error of 6 months. For all tests excepting the easy and hard opposites, a zero score was placed arbitrarily at 5 years as there was no way of determining where it should fall. The method used for finding mental age was that employed in the recent revisions of the Binet-Simon scale.

The results tend to run somewhat lower than do the intelligence quotients derived from the same data by the "more or less" method. The differences between the two series of I. Q. s range from 0 to 9, with an average of 3.4 points. Whatever differences exist are probably due to the system of weighting employed by the "more or less" method, and to the necessity in the use of the "all or none" method for discarding data which falls between the norms of successive ages or beyond the norms of the 13 year group.

### Summary

1. A series of group tests has been arranged that classifies, tentatively, children of grade school age according to mentality. The accuracy of this classification, for the upper age groups, is limited by the shortness of the scale. The value of the series might be increased by the omission of some of the tests, the addition of better tests and the more perfect standardization of whichever of the present tests might be retained.

2. A scale based upon the results secured with this series of tests has been constructed of units which, under certain assumptions, may be regarded as absolute intelligence units. As the method is not limited in its use to group tests, a similar scale could be constructed of absolute intelligence units based upon the results secured by giving the same tests, or better ones of the same general type, to individuals. Such a scale would be free from the objections attaching to all results based upon group tests.

3. A system of weighting has evolved in connection with the method used to reduce scores to units of absolute intelligence. This system of weighting has its origin in the discriminative values of the various tests which, in turn, are based upon the quotient derived from the division of the difference between the norms of two successive age groups by the average of the standard deviations from those norms.

4. The results obtained by the use of the absolute intelligence scale show a positive correlation with chronological age, school grade and with results secured by the individual

testing of subjects.

5. A comparison has been made between the results of the "more or less" and of the "all or none" methods of dealing with the same data. Intelligence quotients were derived by each method for 52 cases and the results were compared with those secured by individual tests given according to the Kuhlmann (1917) revision of the Binet-Simon scale.

In 21.1% of the cases the I. Q.s derived by the "all or none" method are nearer to those secured from the individual tests than are the I. Q.s derived by the "more or less" method.

In 63.4% of the cases the I. Q.s derived by the "more or less" method are nearer the individual I. Q.s than are those derived from the "all or none" method.

In 19.2 % of the cases the difference between the two methods of deriving the I. Q.s results in a change in classification.<sup>1</sup>

The I. Q. s derived by the "more or less" method would appear to be somewhat more exact, but whether they are enough so to be worth the additional step necessary to secure them can be determined only by further investigation.

<sup>1</sup>Terman, 1916, *The Measurement of Intelligence*, p. 79.



### Subjects for Further Research

1. Correlation of intelligence quotients derived from these tests by the "more or less method" with intelligence quotients based on individual tests of the "all or none" type, e. g. Kahlmann's 1917 Revision of the Binet-Simon Scale.

2. Standardization of easy and hard opposites tests on the basis of frequency tables worked out for child subjects as King and Gold have worked theirs out for educated adults.

3. Standardization of a test for associative memory which shall exhibit greater discriminative capacity than the immediate memory tests now in use.

4. Securing data from a larger number of cases in order to note the effect upon the separate and combined curves, and to secure more accurate norms.

5. Replacing less valuable tests with better, i. e. with more discriminative ones, especially with such as involve reasoning ability without stressing speed or language ability.

6. Extension of scale to include higher age groups.

## Appendix

## Material, Procedure and Method of Scoring Tests

Subjects were told that they were going to be asked to do some things that were not hard, but that they would have to listen carefully, and to think, and to do exactly what they were asked to do. They were told also that the papers would be placed on their desks face downward and were not to be touched until the signal was given as any paper turned over before the proper time would be worthless and would have to be thrown away. Detailed directions were given with regard to placing the name, age, grade, etc. on the back of each test sheet.

Materials were handled by the teacher and experimenter for the younger groups but the various monitorial systems of the upper grades were made use of in distributing and collecting them. There was no handling of materials by the experimenter while a test was in progress nor was there any communication between the experimenter and teacher. No questions on the part of the subjects were noticed after a test was once begun. All sources of distraction were eliminated as much as possible.

## Digits

## Material:

The reverse side of the easy opposites test sheet was used for this test in order to save the time which would have been necessary to handle an extra set of papers.

Well sharpened pencil for each subject.

List of digits for the use of the experimenter.

## First Series

6

85

374

2947

35871

491572

2749385

69413827

294758631

## Second Series

2

37

581

6135

92736

516283

6195847

38159462

492875138

For first grade subjects it was necessary to ask the teacher to write the digits on the blackboard at the front of the room and to have the children name them as she wrote. The digits were left on the board during the giving of the test. Ample time was allowed for writing, especially in the first and second grades.

## Formula:

"I am going to read you some numbers and when I am through I want you to write them on your paper just the way I read them. If I said '5' you would write \_\_\_?" Responses were written on the blackboard as subjects gave them. "If I said '4,2' you would write \_\_\_? If I said '1,5,3' you would write \_\_\_?"

"While I am reading the numbers you will hold your pencils up straight like this, with your elbow on your desk. When I am through reading, you will write the numbers and as soon as you have finished writing you will put your pencils up again to show that you are ready for the next."

"Now listen carefully, so that you will be able to write

the numbers just the way I read them. Pencils up. Ready." The signal, "ready", was given about one second before each group was presented. The digits were read at the rate of one per second. Each group was read but once.

**Scoring:**

Correct digits in the correct order.....right.

Correct digits with only one out of place..... $\frac{1}{2}$  right.

Correct digits with more than one out of place.....wrong.

Any error in digits.....wrong.

Right responses were gathered together without regard to the size of the groups of digits and the number of "rights" plus the half credits earned gave the score for each series. For instance, a subject who got the first, second and third groups right, the fourth wrong, the fifth right and the sixth  $\frac{1}{2}$  right, would make a score of  $4\frac{1}{2}$ . The average for the two series was recorded as the final score.

With young subjects inaccuracies in forming the digits were not regarded as errors in cases where the intention was unmistakable, as in a 7 made with with the top toward the right.

**Easy Opposites.**

**Material:**

Printed lists of 20 stimulus words with space for writing responses.

Pencils.

Stop watch.

List of standard responses for scoring.

**Procedure:**

Material was distributed face down on the desks of subjects, the top of the sheet being towards the top of the desk in every case.

Formula:

"This paper has some words on it. I want to see how quickly you can write a word after each of these words that means just the opposite. If the word on the paper were "rich" you would write \_\_\_\_? If the word on the paper were 'black' you would write \_\_\_\_? and so on. The opposite was written after the stimulus word on the blackboard as it was given by the group. If some subjects seemed not to understand the task, it was further illustrated with "up" and "go". "When I say 'ready' take you pencil in the hand you write with, and take hold of the corner of the paper with the other hand. When I say 'turn', turn your paper over, begin to work as fast as you can and keep on until I say 'stop'. The minute I say 'stop' put your pencils down.

Thirty seconds were allowed for the grades 4A to 8A inclusive; sixty seconds, for those below 4A.

Scoring:

The score consisted in the number of correct responses, which were limited to those given by King and Gold. They are given below with their frequencies. Spelling was disregarded wherever the intention of the subject was unmistakable.

<u>Stimulus</u>	<u>Frequency</u>	<u>Response</u>
high	98	low
winter	100	summer
brother	99	sister

<u>Stimulus</u>	<u>Frequency</u>	<u>Response</u>
white	98	black
yes	100	no
new	99	old
come	98	go
asleep	100	awake
north	99	south
sour	98	sweet
dirty	100	clean
dead	99	(alive living
rich	98	poor
wet	100	dry
day	99	night
good	98	bad
east	100	west
empty	98	full
young	100	old
war	100	peace

#### Hard Opposites

#### Material

Printed lists of 20 stimulus words with space for writing responses.

Pencils

Stop watch

List of standard responses for scoring.

#### Procedure:

Same as in the easy opposites test.

Time:

Sixty seconds.

Formula:

"This is another paper with words printed on it. You are to do just as you did before, writing after each of these words a word that means just the opposite. When I say 'ready' -----etc." Continue as in easy opposites.

Scoring:

Same as in easy opposites test.

<u>Stimulus</u>	<u>Frequency</u>	<u>Response</u>
above	94	below under beneath
backwards	94	forwards
beginning	90.5	end ending
broad	93	narrow
buy	94	sell
cheap	93.5	expensive dear costly
forget	94	remember
here	94.5	there elsewhere
hot	92	cold
in	93.5	out
land	90	water sea
late	93	early
left	91.5	right
love	95	hate (hatred)

## VII

<u>Stimulus</u>	<u>Frequency</u>	<u>Response</u>
many	93	few
near	93	far
outside	94.5	inside
quick	95	slow
rough	93	smooth gentle level
something	95	nothing

## Substitution

**Material:**

Woodworth and Wells Substitution Blanks

Pencils.

Stop watch.

**Procedure:**

Material was distributed face down on desks of subjects, with the top of the test sheet towards the top of the desk. The five forms at the top of the sheet were sketched roughly on the blackboard but the numbers were omitted. Below them, eight forms were drawn to be used in illustrating how the filling in was to be done. (Cross, triangle, star, circle, square, star, cross, circle.)

**Time:**

100 seconds.

**Formula:**

"At the top of this sheet of paper (hold up sample) is a row of forms like this (point to forms on board). In each one of these (point to paper again) a number is written -- a different



number in each form."

"Below are a lot of these forms printed in lines. You are to write in each of these (point) the number that is in that same form in the row at the top (point). That is, you are to write in every star you come to the same number that is in the star at the top of the paper, and write in every circle the number that is in the circle at the top of the paper: in every square the number that is in the square at the top: in every one of these (point to cross) the number that is in the one like it, and in every triangle (point) the same number that is in the triangle at the top of the paper. Go straight across the paper this way (point) just as you do when you read and put the right number in each form that you come to."

"When I say 'ready' -----etc." Continue as in the easy opposites test.

"Now, remember, you are to put the same number in each form that is in the same form at the top of the page. Suppose this were the first row (point) what number would you put in here? (Sample row was filled in at dictation of subjects.) Don't skip around, but work straight across the line."

"Ready. Turn."

Scoring:

Scoring consists simply in counting the number of correct substitutions.

Word Building

I

Material:

Test sheets with e a i r l p printed at the top.

Pencils.

Stop watch.

Procedure:

Material was distributed face down on desks of subjects.

Time:

5 minutes.

Formula:

"This paper has some letters printed on it. I want to see how many words you can make from these letters in five minutes. The words must be real ("English" added for grades above the third) words -- words that mean something. You are to use only these letters and you are not to use the same letter more than once in the same word. Of course you do not have to use all the letters in any word. Suppose the letters were e,n,p,a (write on board). What word could you make? As soon as you think of a word, stand! (As soon as entire room was standing, subjects gave words and the group discussed right and wrong illustrations.)

"When I say 'ready' -----etc." Continue instructions as in easy opposites test.

"Now remember, you are going to make just as many words as you can, but the only letters you are going to use are those that are printed at the top of the page."

"Ready. Turn."

Scoring:

See Whipple's manual? Score consists in the number of words correctly formed.

\*Whipple, Guy M., 1915, Manual of Mental and Physical Tests.

## II

**Material:**

Test sheets with a e o b m t printed at the top.

Pencils.

Stop watch.

**Procedure:**

Same as above.

**Time:**

5 minutes.

**Formula:**

"Here is another paper with letters printed on it. You are to see how many words you can make from these letters in five minutes. You will work just as you did before, but the letters at the top of the page are different, and so of course the words you make will be different.

**Scoring:**

Same as in Word Building I. For the final score the results of I and II were averaged.

**Cancellation****Material:**

Test sheets of pisd type (capitals).

Pencils.

Stop watch.

**Procedure:**

Test sheets were distributed face down on desks of subjects with special care that the top of the sheet was towards the top of the desk.

Time:

100 seconds.

Formula:

"Here is a sheet of paper with letters on it all mixed up (hold up sample). I am going to see how fast you can cross out all the Hs and Ms and Ts by drawing a line like this through each one. (Illustrate on blackboard with the three letters.) Cross out all three letters as you go along and don't go back to cross out any that you may have skipped. Go just as fast as you can without skipping any."

"When I say 'ready' -----etc." Continue as in previous tests.

"What letters are you going to cross out?"

"Ready. Turn."

Scoring:

Score consists in the number of correct letters cancelled.

### Anagrams

Materials:

Test sheets of anagrams with space for writing responses.

Pencil.

Stop watch.

Procedure:

Material was distributed face down on desks of subjects.

Time:

5 minutes.

Formula:

On this paper are some words with the letters all mixed

up. You are to change the letters around so that they will make real ("English" inserted for groups above third grade) words -- words that mean something. The words are easy ones, that you know and use every day."

"If the letters were o,d (write letters on blackboard), what word could you make? If the letters were c,t,a, what word could you make? Don't tell but stand up as soon as you think you know." Wait until whole group is standing. "What could you make with n,m,a? Do you understand? You must use all the letters but you mustn't put in any that aren't there?"

"If you come to a word that you can't get, go on to the next word and come back to the hard word later."

"When I say 'ready' -----etc." Continue instructions as in previous tests.

#### Scoring:

Scoring consisted in counting the number of correct solutions.

### Language Completion Test

#### Material:

Trabue Practice Sheets.

Trabue Language Scale B.

Copy of the Trabue manual for scoring responses.

Pencils.

Stop watch.

#### Procedure:

Practice sheets were distributed face up on desks of subjects. Subjects were instructed to read the sentences printed on

\*Trabue, M. R., 1916, Completion Test Language Scales.

the sheets and, as soon as they thought they knew the words that had been left out, to write them in. The fact that just one word must be written in each space and that each sentence must make a sensible statement was stressed. The sheets were collected as soon as they were correctly filled in, but every subject was required to make the correct completions before his sheet was accepted.

Test sheets were then distributed, face down.

Time:

7 minutes.

Formula:

"This paper has some sentences on it with words left out, something like those on the little bit of paper that you have just had. You are to work just as you did before, writing just one word in each blank space and always choosing the word that makes the most sensible statement -- that makes the sentence say something."

"The first thing you will do when you turn your papers over will be to write your name here (show on sample sheet) in the space that is left for it. Then go right to work as hard as you can."

"When I say 'ready' -----etc." Continue as in previous tests.

"Remember, there is to be just one word in each space, and you are always to choose the word that makes the sentence say something."

"Ready. Turn."

Scoring:

Scoring was done in accordance with the Trabue instructions

XIV

excepting that no responses were accepted but those actually listed in the manual.

II

Material:

Trabue Language Scale L.

Pencils.

Stop watch.

Procedure:

Test sheets were distributed face down on desks of subjects.

Time:

5 minutes.

Formula:

"This paper has some more sentences on it with words left out. You are to work just as you did before, remembering always to write just one word in each blank space and to choose the word that makes the most sensible statement."

"Ready. Turn."

Scoring:

Same as for Part I. For the final score the sum of the scores for the two sheets were taken.

Comprehension

I

Material:

Test sheets on each of which four squares and four circles were printed with a dot at the center, top, bottom, right and left of each.

## Pencils.

12 inch cardboard square and a circle 12 inches in diameter to enable the experimenter to draw quickly and accurately the forms needed in the practice test.

## Procedure:

Material was distributed face down on the desks of the subjects. Two squares were drawn on the blackboard.

## Formula:

"Do you see this square? This is the center of the square. (Point to dot) What is it? This is the upper right corner. (Point) What is it? This is the lower left corner of the square. (Point) What is it? This is the middle of the left side of the square. (Point) What is it? This is the middle of the lower side of the square. (Point) What is it?"

"I want a line drawn from the center of the square to the upper left corner. (Pause and repeat last phrase) As soon as you think you know how to draw it, stand." As soon as the majority of group was standing, a subject was asked to draw the line.

"Now I want someone to draw a line from the center of the square to the middle of the left side." When group was standing, another subject was allowed to draw the line.

"On your paper are four squares. On each square you are to draw a line just as I ask you to. Ready. Turn. At the top of your paper make a mark like this  $\dagger$  {draw on blackboard). On the first square draw a line from the center of the square to the lower right corner. (Stress and repeat last phrase)"

"On this second square draw a line from the center to the



square to the middle of the upper side. (Stress and repeat last phrase.)"

"On this third square draw a line from the center of the square to the upper right corner. (Repeat last phrase.)"

"On this fourth square draw a line from the center of the square to the middle of the lower right side. (Repeat last phrase.)"

"Turn your papers over".

(Draw two circles with dots at the centers and extreme upper, lower, right and left points.)

"Do you see this dot at the center of this circle? I want someone to draw a square on this circle so that the upper left corner of the square will be at the center of the circle. (Repeat 'so that ----- circle')"

When a large proportion of the group is standing let a subject draw the square and point out to the rest of the group why it is right.

"On this circle I want a square drawn so that the middle of the right side of the square will be at the center of the circle. (Repeat 'so that ----- circle'.)"

Proceed as before.

"Ready. Turn."

"Be sure that your paper is on your desk so that this (+) mark is at the top. On this first circle (point) you are to draw a square so that the lower left corner of the square will be at the center of the circle. Repeat 'so that ----- circle'.)"

"On this second circle you are to draw a square so that the middle of the upper side will be at the center of the circle. (Repeat 'so that ----- circle'.)"

"On this third circle you are to draw a square so that

the upper right corner will be at the center of the circle. (Repeat from 'so that ----- circle'.)"

"On this last circle draw a square so that the middle of the lower side of the square will be at the center of the circle. (Repeat from 'so that'.)"

**Scoring:**

Scoring consisted in counting the number of lines and squares correctly drawn.

II

**Material:**

Test sheets on which five squares similar to those used in Comprehension I were printed in a vertical row near the right edge of the page. To the left of each square directions taken from the Kuhlmann 1917 Revision were printed.

Test cards E and S of the Kuhlmann Comprehension Test\*

Pencils.

Stop watch.

**Procedure:**

Material was distributed face down on the desks of the subjects. Two squares were drawn on blackboard with the aid of cardboard pattern used in Comprehension I. Teacher was asked to copy directions from Trial Card E and Trial Card S of the 13 year Kuhlmann Comprehension Test.

**Time:**

5 minutes.

**Formula:**

"On each of the squares on this paper you are to draw a

\*Kuhlmann, F., 1917, Measurement of Mental Development. p 79.

triangle according to the directions printed beside it." Have subjects draw various kinds of triangles on the blackboard. "We will try it first with these directions on the board." Proceed according to the instructions for Trial Card E for the individual test, but calling upon different subjects to perform the successive steps. Continue in the same fashion with the directions for Trial Card S.

"Now you are to draw the triangles on the squares on your paper, following the directions step by step, just as we have done with these on the board."

"When I say 'ready'-----etc." Continue as in previous tests.

"Ready. Turn."

#### Scoring:

Scoring consisted in counting the number of sides and corners drawn according to directions. No credit was given for a side or corner correctly placed unless the figure of which it formed a part was a triangle.

For the final score the subjects' scores for the two tests were added.