SOME OF THE MISUSES AND MISINTERPRETATIONS OF INTELLIGENCE TESTS

SOME OF THE MISUSES AND MISINTERPRETATIONS OF INTELLIGENCE TESTS

By

LEONARD LEROY NEAL

Bachelor of Science

Oklahoma Agricultural and Mechanical College

Stillwater, Oklahoma

1935

Submitted to the Department of Education

Oklahoma Agricultural and Mechanical College

In Partial Fulfillment of the Requirements

For the Degree of

MASTER OF SCIENCE

1939

APPROVED:

W. H. Echols
In Charge of Thesis

N. Cougon
Dean of School of Education

D. G. McIntosh
Dean of Graduate School

PREFACE

The writer wishes to thank Dr. Chauncey, Prof. Lackey, and Prof. Echols of the School of Education of the Oklahoma Agricultural and Mechanical College for their helpful suggestions and criticisms. The writer is especially indebted to Prof. Echols for suggesting that he read "Determinism in Education", for the book helped him to decide the subject of his thesis.

L. L. N.

TABLE OF CONTENTS

SOME OF THE MISUSES AND MISINTERPRETATIONS OF INTELLIGENCE TESTS

CHAPTER I

THE OBJECTIVES

It is not the purpose of this thesis to weigh the negative results of intelligence tests against the positive and then draw a conclusion as to the net results. One thesis is not large enough to cover a field so broad; it would probably take several theses to set forth, prove, and sum up all of the good points, then the bad points, and then, if such a thing were possible, subtract one from the other.

This thesis assumes that intelligence tests have benefits, inherent benefits, however few or great they may be, for education. Thus, taking for granted that intelligence tests are beneficial to education, it is the purpose of this thesis to find and point out some of the factors as expressed mostly by literature and teachers that pertain to the misuse and misinterpretation of intelligence tests, with the hope that if intelligence tests are used in educational setups of the future they will have less misuses and misinterpretations. Though a suggestion or two may be made, it is not the purpose of this thesis to present methods or ways of eradicating any of the harmful outcomes of intelligence tests, but to find some of the misuses, with the hope that educators, after knowing what they are, will work out their own ways of avoiding them.

Naturally then, this paper is primarily interested in finding some of the negative uses of results of intelligence tests; so that they may be avoided, thereby making the intelligence tests more valuable when and where they are used.

Though the general purpose of the thesis is to point out some of the misuses and misinterpretations of intelligence tests, the more specific aims are to see:

I. If the I. Q. is constant

II. If authorities agree as to what is measured by intelligence tests

III. If authorities think that intelligence tests are accurate measurements of future success ~~in school or life~~

IV. If ~~teachers and~~ authorities believe:

1. That environment affects intelligence as measured by the present so-called intelligence tests, that intelligence tests do not measure only native ability
2. That there are several different environmental factors that affect the results of intelligence tests
3. That sometimes the tests are not given correctly, and that sometimes they are not graded correctly
4. That intelligence tests discriminate against some people, such as:
   (1) Socially minded people
   (2) Mechanically minded people

(3) People with an English handicap

(4) People with physical defects

(5) People who think slowly but deliberately

5. That there are detrimental effects of intelligence tests, such as:

   (1) Given to satisfy someone's curiosity

   (2) Stigmatizes children

   (3) Gives some students the "big head"

   (4) Promotes laziness among the slower group

   (5) Promotes discord between patrons and the school

6. That in certain cases other tests may be substituted for intelligence tests

7. That too much emphasis has been put upon the I. Q., but that there should be some kind of an I. Q. testing program

CHAPTER II

SOURCE AND TREATMENT OF DATA

The organization of the material throughout the thesis, with the exception of the first chapter or two and the conclusion, will be based upon a questionnaire. After the tabulated results of each item of the questionnaire are presented and discussed, additional discussion may be drawn from at least three other sources---literature, research, and experience.

Most of the discussion in this thesis will be drawn from the tabulated results of the questionnaire and from literature; however the fundamental discussion of a few items will be taken either from experience or research.

The writer knows that personal experience by itself may carry very little weight in proving a point but it may be presented in this thesis merely to corroborate the findings of the questionnaire, literature, or the research. The research part of this thesis, on the constancy of the I. Q., may concern only one item of the approximately eighty items of the questionnaire, yet the research work on this one item will be such that it could be enlarged and presented as a thesis by itself. The writer would like very much indeed to do this, but he knows that if he were to write a thesis on this one item he would prove only one misuse of intelligence tests; whereas he hopes to present several. As a general rule the discussion from literature, like those from the research and experience, will be given not by itself

to prove something but will be presented with the findings of the questionnaire, with the hope that a corroboration of outcomes from two, three, or more sources will carry greater weight than if all the discussion, though it might be greater in amount, could come from one of the four sources to be drawn from.

One of the four or more principal sources from which information was taken for this thesis is the questionnaire to be presented in its entirety toward the end of this chapter. This questionnaire was passed among the students in four different advanced classes in the school of education at the Oklahoma A. and M. College during the summer of 1938. Two of the classes were senior classes, and two were graduate. Prof. Echols taught three---History of Education 423.1, with 53 enrollees, and 423.2, with 27 enrollees, and Philosophy 592, with 13 enrollees, and Dr. Hill, visiting summer lecturer from Harvard University, taught one, with approximately 160 enrollees. Dr. Hill's class returned 75 questionnaires and Prof. Echols' three classes returned 79, making a total of 154.

For more than one reason the information gained from this questionnaire should be more reliable than that gained from the usual questionnaire. The blanks were handed out and taken up by the writer. He was in front of the classroom when they were being filled out, and answered questions that were asked about the instructions, about what was wanted on the blanks. The questionnaire was given only to people

interested in teaching. In Prof. Echols' two classes of History of Education were a few inexperienced teachers but many of them were experienced, some grade teachers, some high school. High school teachers predominated in Prof. Echols' Philosophy 592, though there was a superintendent or two. Administrators predominated in Dr. Hill's class---heads of departments, principals, superintendents, and college instructors. The questionnaire was filled out by responsible people in responsible positions, by representatives from schools all over the state of Oklahoma, including some half dozen colleges, as well as by a few from schools outside of the state of Oklahoma, from as far away as Minnesota.

One member of the college faculty, not a member of my thesis committee but an authority on tests, thought that, perhaps, the questionnaire should have some positive items, items that would require answers favorable to intelligence tests. So the writer changed it to meet his approval, putting in at random 14 positive items with the 15 negative ones under III-1, and put only that part of the original questionnaire back in Prof. Echols' three classes just a day or two after the entire questionnaire had been presented. When the new blank was circulated three or four students who answered the first time were absent and two or three students that were absent the first time were present the second time, with the result that 78 blanks were returned the second time as compared with 79 the first.

III-1 of the questionnaire is presented here with the

results of both circulations. It is self explanatory; and because the positive items changed the answers or results very little, about as much one way as the other, only the original III-1 will be considered in this thesis.

Questionnaire on Intelligence Tests

You do not have to give your name and position, but please give the name of your school:

III. Check the following things that you think may apply to intelligence tests or to the results of intelligence testing programs:

1st	2nd		
63	58	(1)	Not an accurate measurement
	20	(2)	Fair to all students
	30	(3)	Makes teaching easier
	8	(4)	Measures accurately
22	20	(5)	An educational fad
19	20	(6)	Given to satisfy someone's curiosity
9	3	(7)	Out of date
17	19	(8)	Stigmatizes children
	25	(9)	Gives the brighter students a better opportunity
	28	(10)	Beneficial to slower group
	22	(11)	Patrons understand their children better
32	42	(12)	Gives some students the "big head"
	8	(13)	Involves few if any errors in scoring
	16	(14)	Involves few if any errors in administration
10	14	(15)	Promotes laziness among slower group
	43	(16)	Solves some discipline problems
28	20	(17)	Promotes discord between patrons and school

23 29 (18) Causes trouble among children in the same family

38 23 (19) Not given correctly

24 15 (20) Not graded correctly

33 (21) A better basis for the grouping of students than the grades or marks which they make

48 (22) Teachers understand their students better

9 (23) A better guide for promotion than teacher's examination

32 (24) An aid in choosing one's vocation

7 7 (25) Wastes too much of the student's time

30 (26) A means of selecting students who should receive training in institutions of higher learning

1 3 (27) Too much work on the teacher

11 14 (28) Expense of testing too great

28 21 (29) Discrimination against some of the students

If you wish to add to this list, do so below:

_____ _____

The entire original questionnaire, with its some eighty items, follows:

QUESTIONNAIRE ON INTELLIGENCE TESTS

I. You need not give your name and position, but please give below the name of your school:

1. (Your name)_____

2. (Name of School)_____

3. (Position)_____

II.

1. If you think culture or environment affects one's intelligence, check any one or all of the following things that you think might affect one's I. Q.--- keep it low, raise or lower it:

___ (1) Time in school ___ (4) Home life

___ (2) Type of school ___ (5) Church

___ (3) Type of teacher ___ (6) Organization

___ (7) Test training---number of tests, etc.

___ (8) Place reared---city, country, etc.

___ (9) Reading habits

___(10) Habitual rate of work

___(11) Playmates

___(12) Physical work---kind, etc.

If you wish to suggest any other factors that might change the I. Q., do so on the following blanks:

(13)_____ (16)_____

(14)_____ (17)_____

(15)_____ (18)_____

2. Do you think that the so-called intelligence tests measure only native ability, absolutely independent of acquired or environmental ability?
(Yes or no)_____

3. If you think that the so-called I. Q. is influenced by acquired ability or achievement, approximately what per cent of the I. Q. would you say is probably the result of schooling?_____

III.

1. Check the following that you think are detrimental influences or results of I. Q. tests in the school:

____ (1) Not an accurate measurement

____ (2) An educational fad

____ (3) Given to satisfy someone's curiosity

____ (4) Out of date

____ (5) Stigmatizes children

____ (6) Gives some students the "big head"

____ (7) Promotes laziness among slower group

____ (8) Promotes discord between patrons and the school

____ (9) Causes trouble among children in the same family

____(10) Not given correctly

____(11) Not graded correctly

____(12) Wastes too much of the student's time

____(13) Too much work on the teacher

____(14) Expense of testing too great

____(15) Discrimination against some of the students

If you wish to add to this list, do so below:

_____ _____

_____ _____

2. Have teachers put too much emphasis upon the I. Q. in methods of teaching, grouping, etc? (Yes or no)_____

3. Generally speaking, do you think I. Q. testing programs such as we have now should be done away with in the school? (Yes or no)_____

4. Do you believe in some kind of an I. Q. testing program? (Yes or no)_____

5. If I. Q. tests are given at all, should they be more or less limited just to problem children? (Yes or no)_____

6. Does your school have a record of the I. Q.'s of its students? (Yes or no)_____

7. Should all elementary and secondary schools give I. Q. tests to all the children? (Yes or no)_____

8. Do you believe in classifying or grouping students according to their M. A.'s? (Yes or no)_____

IV.

1. Does your school have the money and the time to give several intelligence tests, at least two or three, to every student? (Yes or no)_____

2. Does your school have the money and time to buy and give both intelligence tests and achievement tests? (Yes or no)_____

3. Does a low score on an achievement test stigmatize a child as much as a low score on an intelligence test? (Yes or no)_____

V. You need not answer the following six questions unless intelligence tests have been given some time or other in your school:

1. Do the pupils in your school know what they made on their intelligence tests? (Yes or no)_____

2. Are your students grouped according to their M. A.'s? (Yes or no)_____

3. Approximately what per cent of the students objected to being grouped according to their M. A.'s?_____

4. After knowing the I. Q.'s of the different pupils, did you spend more, less, or about the same amount of time on your slow pupils as you did before you knew their I. Q.'s?_____

5. Is enough benefit derived from knowing the I. Q.'s of students to compensate for the expense, the teacher's time, and the children's time spent in an I. Q. testing program? (Yes or no)_____

6. Regardless of all the disadvantages and drawbacks of the average I. Q. testing program, are the advantages and benefits great enough to justify such a program? (Yes or no)_____

VI. You need not read the eight items below unless intelligence tests were given in your school; if so, underline the one or ones who gave the tests:

1. Teacher who taught the students tested

2. Teacher who did not teach the students tested, but was somewhat closely associated with them as a sponsor, coach, homeroom teacher, etc.

3. Teacher who did not teach the students tested and who was in no way closely associated with them in or out of school

4. Superintendent

5. Principal

6. Secretary to principal or superintendent

7. An expert at testing who had no association with the school children other than being hired or secured by the school to administer the tests

8. Others

VII. You need not read the first nine items below if I. Q. tests have not been given in your school; if they have, underline the name or names of those who graded the papers:

1. (1) Students who took the test

   (2) Other students

   (3) Classroom or subject-matter teacher of the students tested

   (4) Teacher who did not teach the students tested, but was sponsor, coach, homeroom teacher, etc.

   (5) Secretary to the principal or to the superintendent

   (6) Principal

   (7) Superintendent

   (8) An expert at testing who had no association with the school other than being hired or secured temporarily for the purpose of giving the tests

   (9) An expert given a more or less permanent position by the school but who had little or no association with the children tested

2. Were the test papers rechecked by any member or members of the school faculty to see if there were any errors in the first scoring? (Yes or no)_____

VIII. If you have personally given intelligence tests---

1. Did you give the students an extra minute or two for good measure? (Yes or no)_____

2. Did you talk much during the test? (Yes or no)_____

3. Was there any noticeable disturbance during the test? (Yes or no)_____

4. Approximately how many explanations did you make?

   _____

5. Approximately how many students did you test?_____

6. Approximately what per cent of the students did not keep busy on the time tests, but looked out of the window or did something to waste time?_____

7. Did you read carefully and understand all of the test directions before you started the test? (Yes or no)_____

IX. If you wish to make any remarks about I. Q. tests, one way or another, do so below:

CHAPTER III

RESULTS OF INVIRONMENT

It is not the purpose of this chapter to belittle or laud intelligence tests according to the findings set forth here concerning the effects of environment upon intelligence, but to admonish against uses and interpretations that might be wrong in the light of the information to be presented.

> That home and school environment exerts an appreciable influence on scores on an intelligence test and that the present tests are coachable to a considerable extent are.....to be freely admitted. Such admissions, however, do not invalidate the concept of intelligence nor do they overthrow the whole procedure and technique of intelligence measurement. They help only to emphasize the importance of judicious use and interpretation of the results of intelligence tests.[1]

Let us see from the questionnaire what school teachers and administrators think about the effects of environment upon intelligence. The following is II-1 of the questionnaire, so arranged that both the actual number and the per cent checking any one item are given to the left:

II.

1. If you think culture or environment affects one's intelligence, check any one or all of the following things that you think might affect one's I. Q.--- keep it low, raise or lower it:

No.	Per Cent	
98	63	(1) Time in school
99	63	(2) Type of school
101	66	(3) Type of teacher

1. H. S. Chen, The Comparative Coachability of Certain Types of Intelligence Tests, p. 1

136	88	(4) Home life
78	50	(5) Church
50	32	(6) Organization
72	47	(7) Test training---number of tests, etc.
90	58	(8) Place reared---city, country, etc.
121	79	(9) Reading habits
69	42	(10) Habitual rate of work
98	63	(11) Playmates
63	41	(12) Physical work---kind, etc.

There were 154 prospective teachers, teachers, and administrators who returned blanks, and all but 11 checked some of the above items. This means that 143 of 154 indicated that environment affects one's intelligence. Disregarding the fact that other items might have been added to the above list which some of the 11 might then have checked, 92.8 per cent of the teachers believed that environment affects intelligence.

If the 11 who failed to check any one of the items under II-1 did not think intelligence is affected by environment, they should have answered "yes" in II-2, the results of which follow:

II.

2. Do you think that the so-called intelligence tests measure only native ability, absolutely independent of acquired or environmental ability? (Yes or no)___

Blank------8
No-------132
Yes-------14

Instead of 11, the questionnaire gives 14, showing a mathematical gain of three over what might have been expected. This gain might be accounted for either by a discrepancy in marking, or by the supposition that some checked II-1 and then wrote "yes" in II-2, believing that environment influences intelligence but that intelligence tests measure only that part of intelligence which is innate.

In II-1 the per cent believing that environment affects intelligence is 92.8; in II-2, disregarding the 8 blanks and subtracting the 14, the ones who think intelligence tests measure only native ability, from 154, a per cent of 90.9 may be derived, who indirectly indicate that they believe environment affects intelligence.

In either one of the above tables less than 10 per cent of the teachers indicated that intelligence tests measure only native ability, absolutely independent of acquired or environmental ability, and 90 per cent or more indicated that environment affects intelligence.

This means then, ~~that if the results of the questionnaire are correct~~, that scores on intelligence tests should not be interpreted as indicative of pure native ability.

The following quotation from Kimball Young's "An Introductory Sociology" ~~backs up the results of the questionnaire:~~ confirms above

> What do the tests test? The assumption of the early testers in this country was that mental tests measured something innate called intelligence; but it is clear to anyone but the most prejudiced that intelligence and the tests of intelligence all reflect the social and cultural milieu out of which they grow.....it

should be clear that until the psychologist can control the factors of language, and past learning, that is culture, he is not in a position to support the hereditary theory.[2]

Dr. Standing, instructor in sociology at the Oklahoma A. and M. College, says that most of our present day sociologists contend that environment affects intelligence. In looking over the first part of the questionnaire used in this thesis, Dr. Standing remarked, "I am glad to see someone taking the sociologist viewpoint."

Much of the literature of a decade or so ago, especially that immediately after the World War does not agree, perhaps, with the findings of the above two tables; nevertheless there are writings, some of them recent and by important educators, which corroborates the opinion of these teachers as expressed in the questionnaire. For example, Newman[3], after studying 50 pairs of identical twins, decides that environment profoundly modifies intelligence and personality.

Last summer I read this in an Oklahoma paper:

> A young child's I. Q. can be changed for the better under favorable environmental conditions, George Stoddard, director of the child welfare research station at Iowa City, told the nation's teachers Tuesday.
>
> .....
>
> Stoddard said another study disclosed that "the illegitimate children of a large sampling of dull and feeble-minded mothers and out-of-work or laboring-class fathers, if placed in good homes in early infancy, will turn out to be bright children as measured by the best tests now available."[4]

---

2. K. Young, An Introductory Sociology, p. 161

3. K. Young, Ibid., p. 74

4. G. Stoddard, "Child's I. Q. Can Be Improved, Research Chief Tells Teachers", Oklahoma City Times, (June 28, 1938)

Dr. Harold M. Skeels gave this information ~~last year~~ before the American Association for the Advancement of Science:

> He studied 147 waifs taken from their own parents and placed in foster homes. Both the fathers and mothers were, for the most part, of low grade intellectually. Nearly 40 per cent of them, it was determined by intelligence tests, ~~have~~ had I. Q.'s below 80, close to the level of feeble-mindedness. Only 13 per cent were slightly superior to the average.
>
> The children were placed in superior homes. Not a one was below normal in intelligence after a period of adjustment. Sixty-five were intellectually superior. Forty-one were on the genius level.
>
> There was no relationship whatsoever between the intelligence of the children and the intelligence of their parents, Dr. Skeels reported.[5]

W. F. Dearborn[6] says:

> The wave of intelligence testing which has swept over American schools has carried with it some debris: among other things the much-discussed notion of a fixed intellectual endowment, with which a child is born, which neither he, his parents, nor his teachers can by taking thought alter, and which the intelligence tests are designed to measure. The considerations presented in the preceding ~~chapters~~ have, I trust, made clear that, on the contrary, what the intelligence tests measure is definitely affected for better or worse, that it is increased or decreased, by what the home and the school, or the parent and the teacher, do for their children and pupils.

Barnes[7] concludes in his thesis, based upon findings in the Stillwater schools, that---

> it would appear from our own studies and from experiments of others, that while superior intelligence is a great endowment, yet, it is not particularly due to heredity and would soon deteriorate if it were not for the influences of environment.

---

5. T. R. Henry, "The Wandering I. Q.", Jou. of the Nat. Edu. Ass., (Feb., 1938), p. 41

6. W. F. Dearborn, Intelligence Tests, p. 134

7. J. H. Barnes, A Statistical Study of Mental Ability and Achievement of Eight and Ninth Grade Pupils of Stillwater, Oklahoma Junior High School, 1924-25, p. 67

Maxwell and Kilzer[8], authors of "High School Administration", say that---

> The uppermost limit to which a trait may possibly develop in a given individual is doubtless determined by heredity, but within the range of zero to that uppermost point the position attained by a given individual is determined by environment.

Dr. Chen[9], of Columbia, concludes that "all intelligence tests are susceptible to the influence of environment", and H. C. Hines[10], author of "A Guide to Educational Measurements", writes that intelligence tests "reflect not only native ability, but home and school training".

Dozens of other educators might be quoted to show that many of our authorities believe that environment influences intelligence. The writer is well aware that much literature has been written in support of the hereditarian theory, and that intelligence tests measure only native endowment; yet the writer, from looking over some 100 books, pamphlets, articles, etc., would say, not from a careful tabulation but from a general impression, that most of the literature for a few years after the World War might support the theory that intelligence tests measure only native ability but that most of the present day literature supports the theory that intelligence tests measure something else too---the effects of environment. Though the amount of influence exerted by

---

8. C. R. Maxwell and L. R. Kilzer, High School Administration, p. 48

9. H. S. Chen, Op. Cit., p. 3

10. H. C. Hines, A Guide to Edu. Measurements, p. 99

each upon an intelligence test score may still be a matter of conjecture, evidently environment does affect intelligence, and intelligence test scores should be interpreted accordingly.

Just as the amount of influence exerted by environment with all of its different factors is not definitely known, just so it is with the different factors that make up that environment; and since a great many factors go into the making of one's environment, some of the factors should have greater environmental effects than others. Though this is true, it is not the purpose of this chapter to give anything like definite information upon the degree of influence exerted by the different factors---it would probably take several theses to do ~~this~~ that, even if such a thing were possible ---that make up the environment of the average person, but merely to see if environment affects intelligence; however, arranging the different items of II-1 in descending order of the number of checks received by the teachers, the items then fall in order of importance of influence, with the most influencial coming first. Such an arrangement follows:

II.

1. If you think culture or environment affects one's intelligence, check any one or all of the following things that you think might affect one's I. Q.--- keep it low, raise or lower it:

	No.	Per Cent	
1.	136	88	(4) Home life
2.	121	79	(9) Reading habits

	No.	Per Cent		
3.	101	66	(3)	Type of teacher
4.	99	63	(2)	Type of school
5.	98	63	(1)	Time in school
5.	98	63	(11)	Playmates
6.	90	58	(8)	Place reared---city, country, etc.
7.	76	50	(5)	Church
8.	72	47	(7)	Test training---number of tests, etc.
9.	69	42	(10)	Habitual rate of work
10.	63	41	(12)	Physical work---kind, etc.
11.	50	32	(6)	Organization

Thus the results indicate that home life is the most potent factor, and that three of the most influencial factors pertain to school---time in school, type of school, type of teacher.

Since the questionnaire shows that over 90 per cent think that intelligence is affected by environment, and that three of the most powerful factors that make up environment are schooling factors, a table on II-3 will be interesting.

II.

3. If you think that the so-called I. Q. is influenced by acquired ability or achievement, approximately what per cent of the I. Q. would you say is probably the result of schooling?_____

No.	Per Cent
2	0

No.	Per Cent
1	1
3	5
1	7
13	10
7	15
1	17
8	20
12	25
6	30
1	33
2	35
6	40
24	50
5	60
1	66
7	70
3	75
1	80
2	90

Blanks---48

Answers--106

Average--36 per cent

Mode-----50 per cent

The above table shows that a few placed the influence of schooling as high as 60, 70, and 90 per cent, but that most of them placed it at 50. The mode is 50 per cent, but the average is 36 per cent.

The results of chart II-3 do not offer anything definite as to the degree of influence exerted, but they do

offer conclusive evidence that school teachers think that schooling is a powerful environmental factor that has a great deal to do with intelligence.

Burt[11], in his correlation studies of intelligence and schooling, estimated that 54 per cent could be traced to schooling; Willard[12] estimated that approximately 50 per cent was the result of schooling.

Whether schooling is responsible for 45, 50, 55, or some other per cent of our intelligence is not so important as to realize that this one powerful intelligence builder is only one of dozens of things that make up our environment, and that when schooling is coupled with the many other factors of environment, the influence of all on our intelligence must be great, a high percentage of the whole, whatever it is, that is responsible for our intelligence.

If the influence of environment is great, then should we not interpret intelligence tests, scores with ~~this~~ that in mind? Then, may we not say, when dealing with a group of intelligence test scores that are above or below normal, that a high percentage of them is the result of environment? Only last year, before the American Association for the Advancement of Science, Dr. Wellman[13] told of a three year study of children sinking from normal to feebleminded after being placed where they did not get individual attention. How

---

11. W. F. Dearborn, Op. Cit., p. 118
12. R. G. Fuller, Fourteen Is Too Early, p. 3
13. T. R. Henry, Loc. Cit.

could native ability account for children who were once normal, sinking to feebleminded? Environment, then, surely must be considered in interpreting such as the above.

CHAPTER IV

NOT AN ACCURATE MEASUREMENT OF WHAT?

All of the findings of the preceding chapter, all of the information concerning the effects of environment upon intelligence, tend to prove that intelligence tests are not accurate measurements if, as some people think, they measure only native ability, to the utter disregard of ability acquired through environment.

Though psychologists and educators may not define intelligence as such, they have been popularly defined as instruments which measure pure native mental capacity; consequently, if table II-2 showed that 92.8 per cent of the questionnaires indicated that intelligence tests do not measure only native ability, absolutely independent of acquired or environmental ability, then, with the popular definition in mind, what results should be expected from III-1?

III.

1. Not an accurate measurement

	No.	Per Cent
Blank	38	24.7
Checked	116	75.3

Thus, in the above table, 75.3 per cent indicate that the tests are not accurate. In Chapter III, table II-1, 92.8 per cent indicate that intelligence tests are not accurate tests of native intelligence to the exclusion of the effects of environment. Perhaps then, to these teachers who answered

the questionnaire, one of the reasons why intelligence tests are not accurate measurements of intelligence is that the tests are supposed to measure only native intelligence when they also measure the effects of environment.

This means then, that if the results of the questionnaire are correct, scores on intelligence tests should not be interpreted as indicative of pure native ability. At least the table shows that three-fourths of the teachers who answered the questionnaire thought, whether correct or not, that intelligence tests were not accurate measurements. This shows what the majority of the teachers thought, whether right or wrong, and if they are representative of teachers in general, then the expert, the college instructor, or the authority on intelligence tests may see his problem: To set aright either 25 per cent of the teachers, or 75 per cent of the teachers.

Instead of a possible 92.8 per cent or more the results show 75.3 per cent, a difference of more than 17. How can this difference be accounted for? Some of it might be the result of haphazard marking; some of the 17 per cent might have thought that the tests were accurate measurements of intelligence even though they were influenced by the results of environment.

It is possible of course, since it is not known what they thought, that the 75.3 per cent who checked the tests as not accurate measurements would have done so even if the questionnaire had stated that intelligence was the result of

both environment and heredity, but it is the writer's opinion that the influence exerted by the popular conception that intelligence tests measure only native ability, plus the power of suggestion that the first part of the questionnaire, based upon such a conception of intelligence, probably caused some of the teachers to check intelligence tests as inaccurate measurements who otherwise might not have so checked them if intelligence had been defined so as to include both native and environmental intelligence.

This means, in the opinion of the writer, that by changing the definition of the thing measured, perhaps, not such an overwhelming number would check intelligence tests as inaccurate measurements.

So, before it can be decided whether they are accurate measurements or not, it must be decided what they measure. Intelligence must be defined.

R. G. Fuller[1] writes in "Fourteen Is Too Early":

> What intelligence is, nobody knows. No two psychologists agree on a definition. But tests have been devised to measure this unknown intelligence. From one point of view it may be said that the tests are an attempt to define intelligence by measuring it. Boring remarks that "the intelligence which is measured by the tests is simply what the intelligence tests measure".
>
> The tests have undergone refinement and improvement, in the direction of a lessened influence of training and schooling on the test results; but the majority of psychologists deny that native intelligence is measured.

Fuller's statement, though it helps to illustrate the fact that psychologists are not agreed as to what intelligence actually is, may be a little radical, but if nobody knows

---

1. R. G. Fuller, Op. Cit., p. 2

what intelligence is, what right would the writer have to say whether intelligence tests are accurate measurements or not, without first defining intelligence?

According to Henmon[2], "Intelligence is intellect plus knowledge", and knowledge depends upon environment.

Corning[3] says, "Fundamentally the authorities all agree that general intelligence is the ability to learn". If this is true, then an intelligence test to be accurate should measure both native and acquired ability.

Ted Brueckner and Ernest Melby[4], authors of "Diagnostic and Remedial Teaching", say that intelligence may be defined as the capacity to learn.

> Thus, say Gilliland and Jordan, some authors do not admit the term "intelligence" in connection with the tests, but say they are measures of "mentality", thus differentiating sharply between "intelligence", as something which may be acquired or developed, and "mentality", which is something native and may not be developed beyond a certain fixed point. Others do not make this distinction, but use the two expressions interchangeably. In fact, no one has yet been able to give a definition of "intelligence" which is universally acceptable, and so there is now a great deal of confusion in the minds of people generally as to what the "intelligence" tests really do measure. Perhaps the best statement that has been made is that they measure to a really marked degree, <u>ability to do school work</u>.[5]

If intelligence is ability to do school work, then the tests should measure the abilities gained by environment as

---

2. W. F. Dearborn, Op. Cit., p. 94
3. H. M. Corning, After Testing---What?, p. 5
4. T. Brueckner and E. Melby, Diagnostic and Remedial Teaching, p. 70
5. A. R. Gilliland and R. H. Jordan, Educational Measurements and the Classroom Teacher, p. 234

well as those by heredity.

Since, as Clinton H. Allen[6] says, "we can not measure intelligence when we can not define it", the writer takes, for various reasons which will be brought out later, and as a basis to work from, the definition that is practical for schoolmen---ability to do general school work.

The findings so far indicate that intelligence tests are not accurate measurements according to the popular definition of intelligence, that they are not accurate measurements of native ability because of the effects of environment. Now, with intelligence defined as ability to do general school work, the findings up to now might not mean that intelligence tests are not accurate measurements; for if intelligence may be defined as ability to do general school work, the old argument of the effects of environment upon intelligence may be avoided.

This does not mean, however, that the overwhelming number who checked intelligence tests as inaccurate would not do so again even with the practical schoolman's definition of the thing measured---ability to do general school work---but, for reasons brought forth a few pages back---the power of suggestion and the popular conception of intelligence---plus a most simple definition of intelligence, the writer is of the opinion that some of the 75.3 per cent might not then check intelligence tests as inaccurate measurements.

In some of the discussion of the last few pages the

---

6. C. H. Allen, Effects of Intellectual Level, p. 6

writer has tried to show why he believes that many of the 75.3 per cent who checked intelligence tests as inaccurate measurements did so because they thought the tests measured the effects of environment; in the following discussion the writer wishes to give other factors for deeming the tests inaccurate measurements. The discussion, though books have been written about the same subjects, will be short, and is offered here, as some of the preceding discussions, not as conclusive evidence, but for information to be used---

1. In interpreting the accuracy of the measuring capacity of intelligence tests
2. For whatever information, whatever that may be, for accounting for factors other than effects of environment that the 75.3 per cent of the ones who answered the questionnaire might have had in mind when they designated the tests not accurate measurements
3. For support of a more practical definition of intelligence

Terman[7] found that the scores of a vocabulary test given to 631 school children correlated exceedingly high, .91, with the intelligence test scores. This ought to show that the tests are extremely verbal, at least the one Terman used.

The above might be proof that vocabulary tests are good intelligence tests, or that intelligence tests are too verbal to measure general intelligence. Now, are they too verbal? Do they measure verbal intelligence or general intelligence?

Dearborn[8] states that the tests are partial to verbal-minded people. If this is true, then it might explain why

---

7. L. M. Terman, The Intelligence of School Children, p. 309
8. W. F. Dearborn, Op. Cit., p. 114

that in college placement tests language students frequently make higher grades than agriculture students.

If the tests are too verbal, whether they are partial to verbal-minded people or not, all of the factors of environment, especially of schooling, would pronounce them not accurate.

Odell[9] says:

> A common method of validating group intelligence tests has been to compare them with an individual scale, usually the Stanford Revision. Very few group tests yield correlations much higher than .75 with this criterion. The same is true of intercorrelations among group tests.

Evidently the different tests do not measure the same thing or there would be a higher correlation.

Whether the tests are accurate measurements or not they do not seem to be accurate measurements of success in school. In 157 investigations summarized by Douglass[10] the correlation between school marks and I. Q. scores ran from .10 to .67, the average being .44. Not very predicative, are they?

Binet[11] found a correlation of about .45 between I. Q. ratings and school achievement; ~~Pressey, .48~~; Terman, .45; Book, .28.

> It is, ~~says Book~~, conceivable that these rather low correlations between school achievement and intelligence may be due to inaccuracies in our methods of measuring both intelligence and school achievement.[12]

---

9. C. W. Odell, Educational Measurements in High School, p. 403
10. W. S. Gray, Tests and Measurements in Higher Edu., p. 178
11. W. F. Book, The Intelligence of High School Seniors, pp. 105, 106
12. W. F. Book, Ibid., p. 106

Speaking of the above, Book continues---

> This seems to indicate the importance for school success of other factors besides mere intelligence..... That this correlation is relatively low indicates without doubt that other factors besides intelligence enter into the making of a highly successful record in school.[13]

> It is further conceivable, writes Book, that a number of special mental factors may serve to enhance a student's school performance, factors which are quite different from general intelligence. One such factor is a good memory. This may be of far reaching value to a pupil in attaining school success, because most of our school work today draws heavily upon a student's sheer ability to retain and to recall. Other mental characteristics not measured by an intelligence test, such as persistence, effort, mental attitude toward school, etc., might also be possessed by a student with only average ability, and may be deficient or totally lacking in another student who has marked intelligence.[14]

Some years ago Dr. Rigg, now at the Oklahoma A. and M. College, and others gave tests to over 10,000 grade children in St. Louis. About 10 years after the tests were given, when the children were in high school, Dr. Rigg checked on the grades or achievement of a graduating class of 153 pupils and found that the pupil who ranked one in the intelligence test for the some 10,100 pupils, ranked 153 in grades---at the bottom of the class. First in intelligence---last in grades. This may be an exceptional case; nevertheless such cases must be taken into consideration in order to avoid some of the misuses of I. Q. tests.

Colvin[15], in his contributions to the Twenty-First Yearbook, writes:

---

13. W. F. Book, Ibid., p. 107

14. W. F. Book, Ibid., p. 108

15. S. S. Colvin, "Principles Underlying the Construction and Use of Intelligence Tests", Twenty-First Yearbook, (1923), p. 38

Will-to-do a task bulks large in the total school performance.

And he admits that intelligence tests do not measure willingness to work, and concludes:

> Finally, the fact has been emphasized that intelligence tests alone are not sufficient to show the probable efficiency of an individual or his success in school or in life, since character as well as intelligence is a vital element in success or failure.[16]

The different studies of Ross, Proctor, King, and Kelley[17] all show that mental tests are not as indicative of future school success as school marks.

Fuller[18] says there is little relation between one's I. Q. score and his success in school or life.

To quote again from Colvin:

> The intelligence rating may be substantially correct, but other factors may weigh heavily in determining a student's success or failure in college. The most important of these are:
>
> 1. The character of the student, particularly his willingness to hold himself down to a strict mental regimen.
> 2. His ideals and purposes.
> 3. His previous educational training, including his study habits.
> 4. His outside distractions, including work, extra-curricular activities and social engagements.
>
> In the light of these facts it may reasonably be concluded that psychological tests, while a valuable aid in determining a student's ability to do college work, can not be relied upon blindly or exclusively. They must be used together with other materials as a basis for diagnosis and prognosis in connection with educational advice and direction in high school and in college.[19]

---

16. S. S. Colvin, Ibid., p. 42
17. L. H. King, Mental and Interest Tests, p. 9
18. R. G. Fuller, Op. Cit., pp. 6, 7
19. G. M. Whipple, "Intelligence Tests in College and Universities", Twenty-First Yearbook, (1923), p. 265

M. R. Trabue[20] says that the interpretation of test results in educational and vocational guidance is largely negative.

Terman[21] has this to say about intelligence tests not being accurate measurements of future success:

> The child is not all intelligence; his fitness to take up the work of a trade is determined partly by such factors as health, industry, attitude toward school work, and regularity of attendance.

Surely the above quotations from literature show that many of the country's prominent educators are of the opinion that I. Q. scores are poor criteria for predicting the kind and amount of success in school or life.

Some people think that intelligence tests are not accurate measurements of intelligence because they measure "speed" intelligence and not "power" intelligence, since the score depends, more or less, upon how many problems solved and not upon how difficult. Colvin[22] says that the tests do not give the slow, but accurate and thoughtful learner full justice. Consequently the writer is of the opinion that some people might work faster, make a higher score, but not be equal to some grave problem which the slow, deliberate thinker might solve.

From the paragraph just finished naturally comes the

---

20. M. R. Trabue, "The Use of Intelligence Tests in High School", Twenty-First Yearbook, (1923), p. 176

21. L. M. Terman, Op. Cit., pp. 299, 300

22. G. M. Whipple, Loc. Cit.

question, Are intelligence tests accurate measurements of the inventive mind? Probably not, for the inventive mind would evidently have power intelligence, but as evidence for this, there is very little. Spencer, Darwin, Newton, and several other famous men have been cited as evidence for the statement just made, because they made low grades in school, but because they made low grades or failed does not necessarily mean they would have scored low on one of the present-day intelligence tests.

Because, for one thing, few if any of our great creative thinkers have taken intelligence tests, there are little data to prove that the inventive mind can not be accurately measured; and, at the time, it is the opinion of the writer that the foregoing reason might work as well the other way to prove that it has not been conclusively demonstrated that the tests can accurately measure the intelligence of the inventive mind. Though the writer has little evidence to offer one way or the other, he would like to ask this question: What test would be an accurate measurement of the inventive mind that does some creative thinking subconsciously? To be sure, there are dozens of other examples, but, if history is correct, the "Kubla Khan" and the Taj Mahal are products of such minds.

Whether or not intelligence tests can measure power intelligence, whether or not they can measure the inventive mind, and even barring the dreamer or subconscious creative thinker, it probably goes without argument that if the tests

are accurate measurements of the abstract inventive mind, they are not accurate measurements of the practical inventive mind---the mechanically inventive mind.

The tests of so-called general intelligence are made almost entirely for ability to deal with ideas instead of ability to work with things; yet Dewey says:

> The simple facts of the case are that in the great majority of human beings the distinctively intellectual interest is not dominant. They have the so-called practical impulse and disposition.[23]

McFarlane[24] found that practical ability does not correlate closely with general intelligence; and---

> Stenquist, reporting a study in which he used tests of "general mechanical intelligence and ability", says that "at least 40 per cent of the pupils from a typical school, who are below average in general abstract intelligence, are above average in the kind of ability required in the four mechanical tests".[25]

Again, some educators, such as Wells[26], McCall, and Thorndike[27], think that intelligence tests fail to test ability to deal with people, fail to measure the social intelligence.

Thus, Fuller[28] says there are different kinds of intelligence. Thorndike[29] says there are three. And who

---

23. R. G. Fuller, Op. Cit., p. 5
24. R. G. Fuller, Loc. Cit.
25. R. G. Fuller, Loc. Cit.
26. R. G. Fuller, Loc. Cit.
27. M. A. McCall, How To Measure in Education, p. 173
28. R. G. Fuller, Loc. Cit.
29. M. A. McCall, Loc. Cit.

knows but that there are more? At least Thorndike mentions the three that have been indirectly suggested through the discussion---abstract, mechanical, and social.

The discussion and quotations of the last few pages have been given to show that the so-called general intelligence tests only measure one of our three or more I. Q.'s, the abstract intelligence. Consequently, this would indicate the name of the tests should be changed, if they are to be considered, even in a broad sense, accurate measurements.

If the so-called general intelligence tests measure only abstract intelligence instead of mechanical or social, or others if there be others, if they measure environmental or acquired intelligence instead of only native intelligence, if they do not measure character and personality, the driving force behind the intelligence, and can not make an accurate prognosis, not as accurate as that of grades, then perhaps it would be better to limit the measuring capacity of the tests to the practical schoolman's definition of what they measure---ability to do general school work.

Even though what intelligence tests measure were limited to the above to avoid inaccuracies of measurement in a broad or general sense, there would still be other things to consider, some of which follow, in the discussion of other inaccuracies, most of which are more or less minor, before it could be said the tests were more than accurate just in a general way.

Although it has been indirectly stated in the preceding pages that intelligence tests are unfair to people who are not verbal-minded, especially to socially minded and mechanically minded people, nothing has been said of the English handicap that foreigners have in taking the tests.

Dozens of examples could be taken from literature to give information on the above topic, but instead, the conclusions of Dr. Rigg[30], of the Oklahoma A. and M. College, in "Some Further Data on the Language Handicap" are cited.

> The day is past when the psychologist can calmly sit down with a stock of "Stanford-Binet blanks", determine the median I. Q.'s for a few children of native American, German, Jewish and Italian descent and proceed without delay to announce to the world the relative intelligence levels of the nationalities represented.

Besides the above, are the tests discriminating to people who become "rattled" on the tests because of knowing that if they fail adverse criticism will be forthcoming? Are they discrim- to people whose tests are not correctly given or graded? Are they discrim- to people who have had an unfavorable environment?

At least, the tests discriminate against the children of whom Dean Tiegs[31] spoke ~~when he said~~:

> no quote
> From 2 to 5 per cent of our pupils have visual, auditory, or motor-coordination difficulties sufficiently serious to interfere with obtaining a valid intelligence test.

---

30. M. G. Rigg, "Some Further Data on the Language Handicap", The Jou. of Edu. Psy., XIX (April, 1928), p. 252

31. E. W. Tiegs, "Breaking Down The I. Q.", Progressive Edu., (Dec., 1936), p. 605

And from the questionnaire---

III-1-(15) Discrimination against some of the students

Forty-six, almost one-third, indicated that the tests discriminated against some students. Evidently they do, for how could a test measure accurately the intelligence of a child who hardly knew the English language, or who had some eye defect? This is more evidence that in some cases the tests can not measure accurately.

It must be considered, too, in some of the discussion to follow, that a few of the supposed inaccuracies are not inherent, not exactly the fault of the test, grading for example.

An item in the questionnaire pertaining to grading is III-(11).

III.

1. Check the following that you think are detrimental influences or results of I. Q. tests in the school:

_____(11) Not graded correctly

No.	Per Cent
41	26.6

About one-fourth indicated the tests were not accurately graded. Not such a high proportion, yet it expresses the opinion of quite a few teachers. Evidently many of them knew, or thought they knew, of tests being misgraded.

The writer taught in one of the largest schools in western Oklahoma a few years ago and knows that when intelligence tests were given in that school, some of the women

teachers tried but gave up as complete failures at trying to accurately score the tests.

The writer taught in another school in western Oklahoma in which the tests were given, and knows that half of the teachers did not finish the scoring of their tests either because, as they said, they were too hard to score, or because it took too much time.

Kelley[32] makes the statements that all papers should be rechecked by an expert, or in a central office under the supervision of an expert, and that a teacher should not score the papers of his own pupils. VII-1 of the questionnaire will give some information on this.

VII.

1. You need not read the first nine items below if I. Q. tests have not been given in your school; if they have, underline the name or names of those who graded the papers:

No.	Per Cent	
2	3.2	(1) Students who took the test
3	4.7	(2) Other students
14	22.2	(3) Classroom or subject-matter teacher of the students tested
7	11.1	(4) Teacher who did not teach the students tested, but was sponsor, coach, homeroom teacher, etc.
9	14.2	(5) Secretary to the principal or to the superintendent
11	17.6	(6) Principal
15	23.8	(7) Superintendent

32. G. M. Whipple, Op. Cit., pp. 35, 47

1 1.6 (8) An expert at testing who had no association with the school other than being hired or secured temporarily for the purpose of giving the tests

1 1.6 (9) An expert given a more or less permanent position by the school but who had little or no association with the children tested

Out of 154 only 36 (63 checks but only 36 papers), less than one-fourth, indicated by telling who scored the papers that the schools in which they taught had given I. Q. tests some time or other. Kelley would not approve of the scoring by the ones mentioned in the first three items, who make up about one-third of the scores. Only two showed that experts checked the papers.

VII-2 gives some more information on grading.

VII.

2. Were the test papers rechecked by any member or members of the school faculty to see if there were any errors in the first scoring? (Yes or no)_____

Yes--------12
No---------13
Blank-----11

From one-third to two-thirds of the papers were not rechecked. This might have given a chance for some misgrading, too.

Though the scorers may be responsible for much of the misgrading, some of the fault may be in the tests; for example, McLaughlin[33], in his thesis at the Oklahoma A. and M.

---

33. J. A. McLaughlin, A Comparative Study of the Reliability and Validity of the Artificial Language Test in the American Council Psychological Examination, 1931 and 1932 Editions, p. 6

College, showed how it was possible "to beat the test" in the language test of the 1931 and 1932 editions of the American Council Psychological Examination.

> Thus the student could get credit for over sixty-five per cent of that test without trying.....this would not be an accurate method of measuring intelligence, but would be to the advantage of the fast, careless worker.

What do we find in the questionnaire about tests not being given correctly?

III.

1. Check the following that you think are detrimental influences or results of I. Q. tests in the school:

_____(10) Not given correctly

No.	Per Cent
73	47.4

Approximately one-half said the tests are not given correctly.

The next table shows who gave the tests in the schools represented by the ones who filled out the questionnaire.

VI. You need not read the eight items below unless intelligence tests were given in your school; if so, underline the one or ones who gave the tests:

No.	
10---	1. Teacher who taught the students tested
12---	2. Teacher who did not teach the students tested, but was somewhat closely associated with them as a sponsor, coach, homeroom teacher, etc.
3---	3. Teacher who did not teach the students tested, and who was in no way closely associated with them in or out of school
14---	4. Superintendent
10---	5. Principal

1--- 6. Secretary to principal or superintendent

2--- 7. An expert at testing who had no association with the school children other than being hired or secured by the school to administer the tests

2--- 8. Others

Only two papers out of the 36, 5.5 per cent, showed that experts had given the tests. Yet, Gilliland and Jordan[34] say other tests may be the work of the classroom teacher but that intelligence tests should be the work of a specialist.

Table VIII has some information on personal administration of the tests. Out of the 154 who answered the questionnaire 45 had given tests; so let us see what they have to say about their giving the tests correctly.

VIII.

4. Approximately how many students did you test?

Five who had given tests did not answer; 40 gave answers that varied from 1 to 3,000, the 3,000 being given by a college instructor. The total tested equals 7,739; the average tested by the 40 equals 193.

VIII.

7. Did you read carefully and understand all of the test directions before you started the test? (Yes or no)_____

Yes--------39
No---------3
Blank------3

Only three designated they did not understand the directions.

---

34. A. R. Gilliland and R. H. Jordan, Op. Cit., p. 39

VIII.

6. Approximately what per cent of the students did not keep busy on the time tests, but looked out of the window or did something to waste time?____

Sixteen gave no answers; 29 gave answers from 0 to 15. The average for the 29 is 5 per cent. This is not very large, is it?

VIII.

3. Was there any noticeable disturbance during the test? (Yes or no)_____

Yes--------1
No--------42
Blank------2

Only one said there was noticeable disturbance.

VIII.

5. Approximately how many explanations did you make?
_____

Fifteen did not answer; 4 answered "few"; 26 gave answers that varied from 0 to 100. The college instructor who gave the 3,000 tests gave the 100 explanations. The total explanations made by the 26 is 190; the average is 7. Since 15 did not answer and four answered "few", it can not be said that each of the 26 that gave explanations gave only 7 explanations in administering 193 tests, but evidently this is not far wrong. This would mean that explanations were made for 3 or 4 per cent of the students taking the tests. To the writer, who has given tests and watched several administered, this seems exceedingly low, but he realizes that the above is an average.

VIII.

2. Did you talk much during the test? (Yes or no)____

No------44
Yes------1

Again, it is only one.

VIII.

1. Did you give the students an extra minute or two for good measure? (Yes or no)_____

No------40
Yes------5

Five teachers, over 12 per cent, admitted giving extra time.

Giving extra time is one form of cheating that Kelley[35] had reference to when he said intelligence tests should be administered by a specialist.

The writer can give at least a dozen examples of cheating done on intelligence tests in college by individuals taking the tests, and some do not pertain to one or two individuals, either. For example, just before the Iowa Placement Tests were given to the entire student body of the Oklahoma A. and M. College in the school year of 1934-35, the writer witnessed a boy, whose name he has forgotten, open a private letter within the post office and produce a copy of the Iowa Placement Tests to prove to two or three onlookers that people in other colleges were sending copies of the test through the mail to their friends at A. and M. who would be taking the tests a few days later.

Some cheating, such as the above, would be hard even for a specialist to stop, but evidently a specialist would avoid

---

35. T. L. Kelley, "Interpretation of Educational Measurements", p. 45

much that a classroom teacher would fail to avoid, at least this is the opinion expressed by T. L. Kelley[36] in "Interpretation of Educational Measurements".

> There are a host of causes, ~~says McCall~~[37], which have the power to produce large or subtle changes in the personality and behavior of the examiner, which behavior may in turn operate to raise or lower the pupils' scores.

Another reason for having good examiners.

Though several things have been discussed in this chapter ~~thesis~~ that affect or might affect the accuracy of intelligence tests, the writer realizes that by no means have all of the possible factors been considered, but thinks that the most important ones have; and in closing this chapter, gives two paragraphs from the end of the first chapter in Kelley's "Interpretation of Educational Measurements".

> These issues strike deep in social life and individual philosophy. We think of the "old" methods and the "old" subjects of the curriculum as being hoary with precedent and prejudice, but the ruts of the test movement are already so deep that there are many who do not see beyond them. We assume that there is a trait---for example, reading---varying from child to child. Let us question this assumption, for it may be a dozen traits erroneously called one. We assume that tests as given by different teachers and at different times have called forth equal or approximately equal effort; we assume a sufficient sensory and motor equipment; we assume that the sampling as drawn out by the test questions constitutes a fair and sufficient sampling of ability. If we can not avoid making these assumptions, we can at least pause long enough to steep our souls in the conviction that they are present and obscure our findings. If the pause is long enough and well spent, we may secure an estimate of the magnitude of the errors introduced. There is a becoming modesty and reserve in the verdict of a tester who has paused this long and to this outcome.

---

36. T. L. Kelley, Loc. Cit.

37. M. A. McCall, Op. Cit., p. 307

Two plus three has so often totaled five, and two times three so commonly yielded six, that we have assumed test scores may with entire propriety be added, subtracted, multiplied, and divided. They seldom can. Test devisers have apparently been quite successful in obtaining test-score units which are substantially equal and can be added and subtracted, but they have failed quite signally in determining reasonable zero points, so that the product or quotient technique rests upon shifting ground. Let us not forget this, and repeatedly ask, "Do I know that the beginning of the scale of measurement is a sound zero point of ability and that I thus may obtain a meaningful quotient?" The very asking of the question has profoundly stirred our mensurative natures, and answering it "No", as we generally must, robs us at once of a very simple method of interpretation, of a very common source of errors in judgment, and of our fellowship with the get-rich-quick variety of mental-test interpreter. It is not to be desired that the quotient technique be completely discarded, but the writer's immediate purpose will have been accomplished if his readers will but think of the height above zero of an average 12-year-old in a dozen mental tests as being comparable to the height above the water of the rail of a rolling ocean liner as measured at twelve different times. This should be---let us hope it is---a concept to make one dizzy, for uncritically to accept any zero point, however derived, as a proper basis for determining quotients is bewildering and mentally loathsome.[38]

---

38. T. L. Kelley, Op. Cit., pp. 16, 17

CHAPTER V

THE CONSTANCY OF THE I. Q.

As all of the findings in Chapter III, The Results of Environment, have weight upon Chapter IV, Not An Accurate Measurement, just so all of the findings in both Chapters III and IV exert influence on this chapter in the discussion of the constancy of the I. Q. As the effects of environment help to make intelligence tests inaccurate measurements of only native ability, just so they exert an influence on the intelligence that may cause it to vary; in other words, if Chapter III brought forth evidence that intelligence is affected by environment, then it brought forth evidence that the I. Q. might vary, for it stands without argument that environment may vary. Again, anything in Chapter IV that brought forth evidence that intelligence tests might not be accurate measurements, brought forth evidence that the I. Q. may vary. All of the environmental factors, and all of the factors that deem intelligence tests not accurate measurements, work to make the I. Q. vary. And, anything in this chapter that proves the I. Q. will vary, will be proof for Chapter III, that intelligence tests are not accurate measurements.

This chapter, though really a part of the preceding one, is being considered separately because both of its importance and because of the amount and kind of research work to be presented. There are many factors that may make intelligence tests inaccurate measurements---but this, perhaps, is

the most important; besides, as far as the writer has been able to tell from literature, the research work to be presented here will cover more cases than many similar research problems or experiments on the constancy of the I. Q.

Before taking up the research work or statistical study, perhaps some information from literature on this subject might help.

Freeman, Holzinger, and Mitchell[1] report in their studies at Chicago of a significant improvement in intelligence for a group of children that were tested before and several years after being placed in foster homes.

Dearborn gives an example of a girl gaining 37 points in intelligence in three years, of another child gaining 22 points in six months.[2]

Dr. Hellen T. Woolley[3] retested children of the Merrell-Palmer Nursery School and found an average increase of 13.6 points in I. Q.

In the Journal of the N. E. A. you may read---

> A child may change from a high-grade moron to a genius in a few years.
>
> Current conceptions of the stability of the I. Q., to which great significance is attached in most school systems, were torn to shreds before the psychological section of the American Association for the Advancement of Science here today by Dr. Beth I. Wellman, professor of psychology at the University of Iowa. . . . . .

---

1. J. H. Barnes, Op. Cit., p. 60
2. W. F. Dearborn, Op. Cit., pp. 126, 208
3. E. L. Marine, The Effects of Familiarity with the Examiner upon Stanford Binet Test Performance, p. 5

> Dr. Wellman told of experiments which make mince-meat out of this idea of a "mystical intelligence". She gave the cases of five children. One had an I. Q. of 89 at three, of 149 at ten, and of 132 at thirteen. Another started at 98, moved up to 167, and fell back to 143 at twelve years of age. Another moved from 98 to 153 in ten years. These were all ordinary children, whose intelligence increased to that of geniuses. A fifth child with an I. Q. of 124 at three had moved up to the supergenius class of 165 at ten. At twelve she leveled back to 154.[4]

Last year Stoddard told the N. E. A. convention that---

> The best tests available showed that for 600 children who attended the pre-schools sponsored in the station, there was an average gain of 20 I. Q. (intelligence measurement) points.[5]

Terman[6] rechecked 435 pupils to find that the middle 50 per cent of all the changes ranged from a loss of 3.3 to a gain of 5.7 in I. Q. This is not a very large change for the middle 50 per cent, but how about the other 50 per cent? Terman[7] partly answers this when he says that about 5 per cent vary as much as 15 points.

> We do not have, says Terman, an infallible measuring scale, and even if we had we should hardly expect the I. Q. .....to maintain perfect constancy.[8]

The following statistical information is presented here to check the constancy of the I. Q. and to compare the results with the deviations of the I. Q. as found by some of the authors just quoted---Stoddard, 20 points; Woolley, 13.6

---

4. T. R. Henry, Loc. Cit.
5. G. Stoddard, Loc. Cit.
6. H. M. Corning, Op. Cit., p. 8
7. L. M. Terman, Op. Cit., p. 300
8. L. M. Terman, Op. Cit., p. 154

points.

The author has found similar studies based on as few as 30 or a hundred pupils but not very many based on more than three or four hundred cases, with the exception of two studies just mentioned---Terman's and Stoddard's. Stoddard's has 600 cases, and Terman's 435. This study has 321 pupils in one school and 98 in another, making a total of 419 different pupils; but this gives an idea of only one-third of the size of the study, for three different I. Q. scores have been used for each pupil. This makes the study include 1,257 I. Q. scores, more than that of any other similar study that the writer has been able to find.

Each of the 419 pupils took the Henmon-Nelson in April, 1936, while in the Sixth; each pupil took the Otis Self-Administering in Nov., 1937, while in the Eight; and each pupil took the Otis Quick Scoring in Nov., 1938, while in the Ninth.

The first table for each school shows the gain and loss of the Otis Self Administering, based upon the Henmon-Nelson; the second shows the gain and loss of the Otis Quick Scoring, based upon the Henmon-Nelson; the third shows the gain and loss of the Otis Quick Scoring, based upon the Otis Self Administering.

The first set of three tables concerns the scores of the 321 pupils in the Cleveland Junior High School of Tulsa, Oklahoma. There were only two pupils in this school who had not taken all these tests.

CLEVELAND SCHOOL

Key to Tables

In the first two columns---

HN----Henmon-Nelson intelligence test scores
OSA---Otis Self Administering intelligence test scores
OQS---Otis Quick Scoring intelligence test scores

In third column, marked (1), are the gains in points of the score in second column over the score in first column.

In fourth column, marked (2), are the losses in points of the score in second column over the score in first column.

In fifth column, marked (3), a zero designates no change in score.

Table I

HN	OSA	(1)	(2)	(3)
92	101	9		
93	94	1		
96	103	7		
130	122		8	
116	116			0
106	102		4	
108	96		12	
104	115	11		
90	96	6		
87	101	14		
67	70	3		
119	110		9	
110	109		1	

Table II

HN	OQS	(1)	(2)	(3)
92	110	18		
93	103	10		
96	99	3		
130	118		12	
116	107		9	
106	112	6		
108	105		3	
104	117	13		
90	97	7		
87	101	14		
67	68	1		
119	111		8	
110	114	4		

Table III

OSA	OQS	(1)	(2)	(3)
101	110	9		
94	103	9		
103	99		4	
122	118		4	
116	107		9	
102	112	10		
96	105	9		
115	117	2		
96	97	1		
101	101			0
70	68		2	
110	111	1		
109	114	5		

Table I

HN	OSA	(1)	(2)	(3)
86	99	13		
113	114	1		
120	121	1		
96	93		3	
97	97			0
94	96	2		
118	121	3		
98	97		1	
99	115	16		
101	109	8		
115	117	2		
104	115	11		
102	95		7	
102	92		10	
121	109		12	
79	62		17	
85	88	3		
96	104	8		
115	113		2	
124	125	1		
84	85	1		
101	100		1	
112	121	9		
98	108	10		
81	90	9		

Table II

HN	OQS	(1)	(2)	(3)
86	98	12		
113	113			0
120	116		4	
96	88		8	
97	106	9		
94	101	7		
118	119	1		
98	103	5		
99	110	11		
101	103	2		
115	123	8		
104	109	5		
102	111	9		
102	104	2		
121	109		12	
79	87	8		
85	94	9		
96	107	11		
115	112		3	
124	126	2		
84	102	18		
101	108	7		
112	118	6		
98	108	10		
81	98	17		

Table III

OSA	OQS	(1)	(2)	(3)
99	98		1	
114	113		1	
121	116		5	
93	88		5	
97	106	9		
96	101	5		
121	119		2	
97	103	6		
115	110		5	
109	103		6	
117	123	6		
115	109		6	
95	111	16		
92	104	12		
109	109			0
62	87	25		
88	94	6		
104	107	3		
113	112		1	
125	126	1		
85	102	17		
100	108	8		
121	118		3	
108	108			0
90	98	8		

Table I

HN	OSA	(1)	(2)	(3)
107	105		2	
77	77			0
104	106	2		
99	89		10	
108	107		1	
117	119	2		
78	90	12		
74	76	2		
112	105		7	
91	81		10	
86	92	6		
86	90	4		
94	83		11	
123	116		7	
88	91	3		
81	91	10		
95	108	13		
84	99	15		
104	100		4	
98	104	6		
86	96	10		
97	106	9		
78	82	4		
108	101		7	
112	118	6		

Table II

HN	OQS	(1)	(2)	(3)
107	109	2		
77	84	7		
104	111	7		
99	96		3	
108	106		2	
117	113		4	
78	94	16		
74	80	6		
112	120	8		
91	80		11	
86	91	5		
86	89	3		
94	93		1	
123	116		7	
88	94	6		
81	92	11		
95	111	16		
84	102	18		
104	113	9		
98	108	10		
86	109	23		
97	91		6	
78	83	5		
108	100		8	
112	118	6		

Table III

OSA	OQS	(1)	(2)	(3)
105	109	4		
77	84	7		
106	111	5		
89	96	7		
107	106		1	
119	113		6	
90	94	4		
76	80	4		
105	120	15		
81	80		1	
92	91		1	
90	89		1	
83	93	10		
116	116			0
91	94	3		
91	92	1		
108	111	3		
99	102	3		
100	113	13		
104	108	4		
96	109	13		
106	91		15	
82	83	1		
101	100		1	
118	118			0

Table I

HN	OSA	(1)	(2)	(3)
96	103	7		
84	73		11	
96	112	16		
116	113		3	
120	116		4	
89	87		2	
91	89		2	
109	111	2		
129	129			0
95	97	2		
91	98	7		
104	106	2		
110	121	11		
103	110	7		
112	109		3	
99	92		7	
112	104		8	
79	82	3		
86	92	6		
72	75	3		
132	120		12	
107	100		7	
105	106	1		
84	85	1		
117	128	11		

Table II

HN	OQ3	(1)	(2)	(3)
96	98	2		
84	82		2	
96	107	11		
116	116			0
120	108		12	
89	93	4		
91	93	2		
109	102		7	
129	130	1		
95	91		4	
91	106	15		
104	101		3	
110	120	10		
103	113	10		
112	114	2		
99	97		2	
112	108		4	
79	88	9		
86	113	27		
72	80	8		
132	120		12	
107	117	10		
105	105			0
84	91	7		
117	121	4		

Table III

OSA	OQ3	(1)	(2)	(3)
103	98		5	
73	82	9		
112	107		5	
113	116	3		
116	108		8	
87	93	6		
89	93	4		
111	102		9	
129	130	1		
97	91		6	
98	106	8		
106	101		5	
121	120		1	
110	113	3		
109	114	5		
92	97	5		
104	108	4		
82	88	6		
92	113	11		
75	80	5		
120	120			0
100	117	17		
106	105		1	
85	91	6		
128	121		7	

Table I					Table II					Table III				
HN	OSA	(1)	(2)	(3)	HN	OQS	(1)	(2)	(3)	OSA	OQS	(1)	(2)	(3)
96	105	9			96	106	10			105	106	1		
101	110	9			101	103	2			110	103		7	
137	112		25		137	119		18		112	119	7		
104	102		2		104	104			0	102	104	2		
80	91	11			80	96	16			91	96	5		
109	116	7			109	115	6			116	115		1	
114	112		2		114	108		6		112	108		4	
90	109	19			90	101	11			109	101		8	
109	117	8			109	117	8			117	117			0
82	78		4		82	86	4			78	86	8		
88	94	6			88	90	2			94	90		4	
124	113		11		124	124			0	113	124	11		
87	94	7			87	101	14			94	101	7		
114	112		2		114	112		2		112	112			0
93	93			0	93	103	10			93	103	10		
92	92			0	92	91		1		92	91		1	
75	91	16			75	96	21			91	96	5		
107	112	5			107	105		2		112	105		7	
114	107		7		114	112		2		107	112	5		
110	108		2		110	113	3			108	113	5		
106	116	10			106	113	7			116	113		3	
111	107		4		111	108		3		107	108	1		
110	112	2			110	101		9		112	101		11	
132	131		1		132	131		1		131	131			0
79	85	6			79	90	11			85	90	5		

Table I

HN	OSA	(1)	(2)	(3)
90	87		3	
70	80	10		
109	112	3		
78	89	11		
94	100	6		
95	95			0
130	113		17	
83	84	1		
95	98	3		
121	113		8	
103	111	8		
106	109	3		
108	109	1		
96	103	7		
85	97	12		
65	66	1		
85	99	14		
108	113	5		
82	91	9		
115	102		13	
96	92		4	
113	109		4	
93	109	16		
83	88	5		
102	86		16	

Table II

HN	OQS	(1)	(2)	(3)
90	100	10		
70	78	8		
109	115	6		
78	99	21		
94	97	3		
95	99	4		
130	120		10	
83	83			0
95	105	10		
121	111		10	
103	108	5		
106	113	7		
108	110	2		
96	105	9		
85	102	17		
65	72	7		
85	104	19		
108	103		5	
82	99	17		
115	109		6	
96	90		6	
113	115	2		
93	105	12		
83	96	13		
102	93		9	

Table III

OSA	OQS	(1)	(2)	(3)
87	100	13		
80	78		2	
112	115	3		
89	99	10		
100	97		3	
95	99	4		
113	120	7		
84	83		1	
98	105	7		
113	111		2	
111	108		3	
109	113	4		
109	110	1		
103	105	2		
97	102	5		
66	72	6		
99	104	5		
113	103		10	
91	99	8		
102	109	7		
92	90		2	
109	115	6		
109	105		4	
88	96	8		
86	93	7		

Table I

HN	OSA	(1)	(2)	(3)
110	107		3	
93	83		10	
68	81	13		
109	120	11		
80	79		1	
129	124		5	
102	102			0
103	115	12		
118	117		1	
107	102		5	
107	101		6	
73	74	1		
99	102	3		
100	100			0
65	65			0
101	112	11		
104	110	6		
99	102	3		
108	113	5		
100	111	11		
75	73		2	
110	116	6		
85	89	4		
105	101		4	
112	109		3	

Table II

HN	OQS	(1)	(2)	(3)
110	85		25	
93	103	10		
68	81	13		
109	111	2		
80	90	10		
129	124		5	
102	112	10		
103	108	5		
118	117		1	
107	110	3		
107	108	1		
73	90	17		
99	112	13		
100	103	3		
65	69	4		
101	112	11		
104	116	12		
99	108	9		
108	114	6		
100	109	9		
75	85	10		
110	116	6		
85	89	4		
105	106	1		
112	102		10	

Table III

OSA	OQS	(1)	(2)	(3)
107	85		22	
83	103	20		
81	81			0
120	111		9	
79	90	11		
124	124			0
102	112	10		
115	108		7	
117	117			0
102	110	8		
101	108	7		
74	90	16		
102	112	10		
100	103	3		
65	69	4		
112	112			0
110	116	6		
102	108	6		
113	114	1		
111	109		2	
73	85	12		
116	116			0
89	89			0
101	106	5		
109	102		7	

Table I

HN	OSA	(1)	(2)	(3)
94	95	1		
100	92		8	
106	110	4		
95	97	2		
86	99	13		
120	115		5	
92	91		1	
103	104	1		
98	100	2		
78	71		7	
109	111	2		
73	73			0
80	98	18		
94	97	3		
96	96			0
83	82		1	
130	119		11	
108	106		2	
75	63		12	
82	86	4		
126	122		4	
106	117	11		
118	115		3	
98	107	9		
88	91	3		

Table II

HN	OQS	(1)	(2)	(3)
94	100	6		
100	102	2		
106	116	10		
95	100	5		
86	96	10		
120	128	8		
92	96	4		
103	108	5		
98	101	3		
78	71		7	
109	113	4		
73	86	13		
80	107	27		
94	108	14		
96	106	10		
83	91	8		
130	114		16	
108	113	5		
75	73		2	
82	97	15		
126	123		3	
106	109	3		
118	106		12	
98	101	3		
88	90	2		

Table III

OSA	OQS	(1)	(2)	(3)
95	100	5		
110	116	6		
92	102	10		
97	100	3		
99	96		3	
115	128	13		
91	96	5		
104	108	4		
100	101	1		
71	71			0
111	113	2		
73	86	13		
98	107	9		
97	108	11		
96	106	10		
82	91	9		
119	114		5	
106	113	7		
63	73	10		
86	97	11		
122	123	1		
117	109		8	
115	106		9	
107	101		6	
91	90		1	

Table I

HN	OSA	(1)	(2)	(3)
106	111	5		
89	104	15		
60	64	4		
113	104		9	
77	59		18	
132	111		21	
84	95	11		
94	109	15		
125	123		2	
92	96	4		
80	86	6		
132	122		10	
111	105		6	
105	113	8		
91	108	17		
88	85		3	
53	57	4		
91	98	7		
105	105			0
103	99		4	
107	129	22		
101	103	2		
99	111	12		
102	104	2		
87	90	3		

Table II

HN	OQS	(1)	(2)	(3)
106	111	5		
89	105	16		
60	79	19		
113	113			0
77	59		18	
132	119		13	
84	101	17		
94	114	20		
125	120		5	
92	100	8		
80	95	15		
132	122		10	
111	108		3	
105	113	8		
91	109	18		
88	90	2		
53	57	4		
91	104	13		
105	105			0
103	100		3	
107	125	18		
101	101			0
99	112	13		
102	100		2	
87	98	11		

Table III

OSA	OQS	(1)	(2)	(3)
111	111			0
104	105	1		
64	79	15		
104	113	9		
59	59			0
111	119	8		
95	101	6		
109	114	5		
123	120		3	
96	100	4		
86	95	9		
122	122			0
105	108	3		
113	113			0
108	109	1		
85	90	5		
57	57			0
98	104	6		
105	105			0
99	100	1		
129	125		4	
103	101		2	
111	112	1		
104	100		4	
90	98	8		

Table I

HN	OSA	(1)	(2)	(3)
93	101	8		
94	104	10		
83	79		4	
100	109	9		
84	88	4		
126	118		8	
76	86	10		
111	99		12	
120	116		4	
93	101	8		
100	106	6		
96	95		1	
131	121		10	
126	119		7	
79	86	7		
73	65		8	
112	124	12		
114	115	1		
118	121	3		
113	109		4	
70	81	11		
76	80	4		
76	76			0
98	105	7		
98	100	2		
136	124		12	

Table II

HN	OQS	(1)	(2)	(3)
93	99	6		
94	101	7		
83	84	1		
100	109	9		
84	88	4		
126	121		5	
76	98	22		
111	106		5	
120	110		10	
93	102	9		
100	116	16		
96	102	6		
131	126		5	
126	116		10	
79	90	11		
73	81	8		
112	123	11		
114	117	3		
118	123	5		
113	87		26	
70	78	8		
76	87	11		
76	86	10		
98	98			0
98	92		6	
136	132		4	

Table III

OSA	OQS	(1)	(2)	(3)
101	99		2	
104	101		3	
79	84	5		
109	109			0
88	88			0
118	121	3		
86	98	12		
99	106	7		
116	110		6	
101	102	1		
106	116	10		
95	102	7		
121	126	5		
119	116		3	
86	90	4		
65	81	16		
124	123		1	
115	117	2		
121	123	2		
109	87		22	
81	78		3	
80	87	7		
76	86	10		
105	98		7	
100	92		8	
124	132	8		

Table I

HN	OSA	(1)	(2)	(3)
97	97			0
107	114	7		
102	108	6		
114	109		5	
104	110	6		
115	111		4	
98	97		1	
134	122		12	
89	90	1		
135	128		7	
92	95	3		
86	90	4		
106	110	4		
92	101	9		
103	99		4	
103	106	3		
115	112		3	
73	79	6		
112	112			0
95	100	5		
78	87	9		
120	136	16		
124	104		20	
107	113	6		
83	96	13		
75	85	10		

Table II

HN	OQS	(1)	(2)	(3)
97	87		10	
107	101		6	
102	115	13		
114	113		1	
104	108	4		
115	116	1		
98	98			0
134	132		2	
89	97	8		
135	135			0
92	108	16		
86	100	14		
106	109	3		
92	103	11		
103	99		4	
103	111	8		
115	107		8	
73	84	11		
112	116	4		
95	106	11		
78	94	16		
120	121	1		
124	107		17	
107	120	13		
83	105	22		
75	94	19		

Table III

OSA	OQS	(1)	(2)	(3)
97	87		10	
114	101		13	
108	115	7		
109	113	4		
110	108		2	
111	116	5		
97	98	1		
122	132	10		
90	97	7		
128	135	7		
95	108	13		
90	100	10		
110	109		1	
101	103	2		
99	99			0
106	111	5		
112	107		5	
79	84	5		
112	116	4		
100	106	6		
87	94	7		
136	121	15		
104	107	3		
113	120	7		
96	105	9		
85	94	9		

Table I

HN	OSA	(1)	(2)	(3)
80	73		7	
77	85	8		
101	102	1		
104	111	7		
78	61		17	
126	125		1	
87	74		13	
125	114		11	
82	81		1	
83	101	18		
117	110		7	
86	81		5	
94	100	6		
102	103	1		
116	113		3	
85	100	15		
120	99		21	
134	114		20	
88	105	17		
133	122		11	
87	98	11		
99	101	2		
90	84		6	
91	90		1	
92	90		2	
99	107	8		

Table II

HN	OQS	(1)	(2)	(3)
80	92	12		
77	91	14		
101	98		3	
104	115	11		
78	82	4		
126	123		3	
87	94	7		
125	113		12	
82	100	18		
83	101	18		
117	114		3	
86	92	6		
94	101	7		
102	106	4		
116	106		10	
85	99	14		
120	110		10	
134	123		11	
88	102	14		
133	132		1	
87	110	23		
99	105	6		
90	91	1		
91	95	4		
92	96	4		
99	111	12		

Table III

OSA	OQS	(1)	(2)	(3)
73	92	19		
85	91	6		
102	98		4	
111	115	4		
61	82	21		
125	123		2	
74	94	20		
114	113		1	
81	100	19		
101	101			0
110	114	4		
81	92	11		
100	101	1		
103	106	3		
113	106		7	
100	99		1	
99	110	11		
114	123	9		
105	102		3	
122	132	10		
98	110	12		
101	105	4		
84	91	7		
90	95	5		
90	96	6		
107	111	4		

Table I

HN	OSA	(1)	(2)	(3)
85	101	16		
144	128		16	
105	107	2		
96	96			0
119	125	6		
96	114	18		
148	124		24	
124	112		12	
115	119	4		
87	98	11		
94	95	1		
79	82	3		
117	109		8	
92	90		2	
91	93	2		
92	94	2		
93	106	13		
87	81		6	
111	109		2	
120	127	7		
118	126	8		
106	120	14		
124	116		8	
106	93		13	
72	79	7		
105	107	2		

Table II

HN	OQS	(1)	(2)	(3)
85	103	18		
144	130		14	
105	105	1		
96	96			0
119	122	3		
96	114	18		
148	137		11	
124	102		22	
115	116	1		
87	105	18		
94	103	9		
79	90	11		
117	101		16	
92	96	4		
91	109	8		
92	98	6		
93	96	3		
87	86		1	
111	115	4		
120	117		3	
118	117		1	
106	118	12		
124	111		13	
106	103		3	
72	83	11		
105	105			0

Table III

OSA	OQS	(1)	(2)	(3)
101	103	2		
128	130	2		
107	106		1	
96	96			0
125	122		3	
114	114			0
124	137	13		
112	102		10	
119	116		3	
98	105	7		
95	103	8		
82	90	8		
109	101		8	
90	96	6		
93	109	16		
94	98	4		
106	96		10	
81	86	5		
109	115	6		
127	117		10	
126	117		9	
120	118		2	
116	111		5	
93	103	10		
79	83	4		
107	105		2	

Table I

HN	OSA	(1)	(2)	(3)
83	71		12	
124	111		13	
91	100	9		
124	116		8	
Total Stu.		179	125	17
Total Pts.		1285	875	0
Mean		7.18	7.0	0

Table II

HN	OQS	(1)	(2)	(3)
83	93	10		
124	111		13	
91	99	8		
124	111		13	
Total Stu.		219	88	14
Total Pts.		1953	651	0
Mean		8.91	7.4	0

Table III

OSA	OQS	(1)	(2)	(3)
71	93	22		
111	111			0
100	99		1	
116	111		5	
Total Stu.		196	96	29
Total Pts.		1373	480	0
Mean		6.95	5.0	0

Comparing scores of the second test with the first, 17 remained the same; comparing scores of the third test with those of the first, 14 remained the same; comparing scores of the third with the second, 29 remained the same. This shows that out of 321 very few remained the same; they either lost or gained.

For the second test, Table I, 125 pupils made an average loss over the scores of the first test of 7 points; in the third test, Table II, 88 pupils made an average loss over the scores of the first test of over 7 points; and in the third test, Table III, 96 pupils made an average loss over the scores of the second test of five points. This means that with each new test there were fewer pupils who made scores lower than their first scores.

In the second test, Table I, 179 pupils made an average gain over the scores of the first test of over 7 points; in the third test, Table II, 219 pupils made an average gain of

almost 9 points over the scores of the first test; in the third test, Table III, 196 pupils made an average gain of almost 7 points over the scores of the second test. The second test scores gained on the first, and the third gained on both the second and first, but gained more on the first than on the second. This means that with each new test there was an increase in the number of pupils who made scores higher than their first test scores.

Not only do the above tables show that the I. Q. changed some, but that it increased. This is made more evident by comparing test three with test one, Table II. By adding the total gained points, 1953, and the total lost points, 651, gives a change of 2604 for 321 pupils. This is an average change of 8.11 points. By subtracting 651 from 1953, the net gain is 1302 points. This is an average net gain of over 4.05 points.

Now, what would account for each pupil gaining over 4 points in intelligence? Could more testing experience? More schooling? Or adolescence? Whatever the cause, evidently it is not native ability.

Though the above results show that the I. Q. was not constant for the Tulsa pupils, it was more constant than the writer expected, for the average change of 8.11 points, though significant, is not as great as the average gain in either the previously mentioned reports of Stoddard or Woolley.

The writer would suggest that a change in environment had a great deal to do with the increase in the I. Q. of the nursery and preschool children of the above studies, for the

children were tested before and after having the influence of good schools. Better environment, especially in the way of schooling, was hardly possible with the pupils of the Cleveland School, for they had the influence of an excellent school environment to begin with.

Cleveland Junior High School is a good school, one of the best and largest in Tulsa. The children who attend it are from stable homes. Knowing this, the writer decided to see how constant the I. Q. was for children in a school where the factors of school and life were more variable.

The scores to follow come from 98 pupils of the Lowell School on the outskirts of Tulsa, where homes are not so stable and life is more variable. There were over 50 pupils in this school who had not taken all three of the tests.

LOWELL SCHOOL

Key to Tables

In the first two columns---

HN----Henmon-Nelson intelligence test scores
OSA---Otis Self Administering intelligence test scores
OQS---Otis Quick Scoring intelligence test scores

In third column, marked (1), are the gains in points of the score in second column over the score in first column.

In fourth column, marked (2), are the losses in points of the score in second column over the score in first column.

In fifth column, marked (3), a zero designates no change in score.

Table I

HN	OSA	(1)	(2)	(3)
71	82	11		
133	118		15	
113	111		2	
87	91	4		
103	114	11		
110	110			0
103	98		5	
80	95	15		
86	80		6	
101	106	5		
100	105	5		
93	102	9		
70	72	2		
103	106	3		
96	103	7		
88	103	15		
63	75	12		
89	110	21		
89	100	11		
72	86	14		
98	96		2	
96	107	11		
85	102	17		
114	118	4		
98	106	8		
105	105			0

Table II

HN	OQS	(1)	(2)	(3)
71	77	6		
133	111		22	
113	115	2		
87	92	5		
103	108	5		
110	112	2		
103	108	5		
80	95	15		
86	94	8		
101	106	5		
100	104	4		
93	108	15		
70	80	10		
103	115	12		
96	85		11	
88	108	20		
63	77	14		
89	105	16		
89	93	4		
72	91	19		
98	101	3		
96	105	9		
85	96	11		
114	114			0
98	110	12		
105	110	5		

Table III

OSA	OQS	(1)	(2)	(3)
82	77		5	
118	111		7	
111	115	4		
91	92	1		
114	108		6	
110	112	2		
98	108	10		
95	95			0
80	94	14		
106	106			0
105	104		1	
102	108	6		
72	80	8		
106	115	9		
103	85		18	
103	108	5		
75	77	2		
110	105		5	
100	93		7	
86	91	5		
96	101	5		
107	105		2	
102	96		6	
118	114		4	
106	110	4		
105	110	5		

Table I

HN	OSA	(1)	(2)	(3)
82	89	7		
89	86		3	
94	99	5		
86	97	11		
114	118	4		
114	119	5		
85	108	23		
94	100	6		
58	63	5		
73	75	2		
102	97		5	
107	109	2		
97	106	9		
99	97		2	
86	100	14		
93	96	3		
81	83	2		
87	85		2	
106	113	7		
73	95	22		
125	127	2		
111	122	11		
138	124		14	
108	117	9		
78	90	12		
104	110	6		

Table II

HN	OQS	(1)	(2)	(3)
82	99	17		
89	97	8		
94	110	16		
86	105	19		
114	124	10		
114	119	5		
85	111	26		
94	104	10		
58	65	7		
73	74	1		
102	92		10	
107	114	7		
97	112	15		
99	105	6		
86	104	18		
93	101	8		
81	97	16		
87	80		7	
106	118	12		
73	94	21		
125	120		5	
111	126	15		
138	132		6	
108	107		1	
78	92	14		
104	107	3		

Table III

OSA	OQS	(1)	(2)	(3)
89	99	10		
86	97	11		
99	110	11		
97	105	8		
118	124	6		
119	119			0
108	111	3		
100	104	4		
63	65	2		
75	74		1	
97	92		5	
109	114	5		
106	112	6		
97	105	8		
100	104	4		
96	101	5		
83	97	14		
85	80		5	
113	118	5		
95	94		1	
127	120		7	
122	126	4		
124	132	8		
117	107		10	
90	92	2		
110	107		3	

Table I

HN	OSA	(1)	(2)	(3)
87	94	7		
118	120	2		
100	101	1		
91	106	5		
108	104		4	
89	102	13		
84	92	8		
92	94	2		
81	97	16		
96	105	9		
87	93	6		
90	102	12		
118	124	6		
72	83	11		
98	94		4	
96	98	2		
106	103		3	
79	85	6		
80	88	8		
86	93	7		
98	96		2	
82	90	8		
131	115		16	
81	96	15		
76	86	10		
75	88	13		

Table II

HN	OQS	(1)	(2)	(3)
87	104	17		
118	120	2		
100	107	7		
91	111	20		
108	106		2	
89	105	16		
84	89	5		
92	106	14		
81	102	21		
96	99	3		
94	87		7	
90	100	10		
118	119	1		
72	92	20		
98	99	1		
96	108	12		
106	105		1	
79	102	23		
80	102	22		
86	99	13		
98	95		3	
82	76		6	
131	117		14	
81	96	15		
76	88	12		
75	97	22		

Table III

OSA	OQS	(1)	(2)	(3)
94	104	10		
120	120			0
101	107	6		
106	111	5		
104	106	2		
102	105	3		
92	89	3		
94	106	12		
97	102	5		
105	99		6	
94	93		1	
102	100		2	
124	119		5	
83	92	9		
94	99	5		
98	108	10		
103	105	2		
85	102	17		
88	102	14		
93	99	6		
96	95		1	
90	76		14	
115	117	2		
96	96			0
86	88	2		
88	97	9		

Table I

HN	OSA	(1)	(2)	(3)
72	83	11		
80	74		6	
88	107	19		
94	105	11		
73	88	15		
75	77	2		
81	107	26		
100	96		4	
76	86	10		
98	99	1		
61	59		2	
74	88	14		
106	116	10		
82	71		11	
96	97	1		
78	85	7		
106	106			0
101	113	12		
111	115	4		
89	102	13		
Total Stu.		77	18	3
Total Pts.		668	115	0
Mean		8.67	6.4	0

Table II

HN	OQS	(1)	(2)	(3)
72	90	18		
80	83	3		
88	106	18		
94	108	14		
73	93	20		
75	96	21		
81	98	17		
100	105	5		
76	95	19		
98	103	5		
61	58		3	
74	96	22		
106	115	9		
82	85	3		
96	106	10		
78	98	20		
106	105		1	
101	114	13		
111	109		2	
89	98	9		
Total Stu.		81	16	1
Total Pts.		943	95	0
Mean		11.6	5.9	0

Table III

OSA	OQS	(1)	(2)	(3)
83	90	7		
74	83	9		
107	106		1	
105	108	3		
88	93	5		
77	96	19		
107	98		9	
96	105	9		
86	95	9		
99	103	4		
59	58		1	
88	96	8		
116	115		1	
71	85	14		
97	106	9		
85	98	13		
106	105		1	
113	114	1		
115	109		6	
102	98		4	
Total Stu.		63	30	5
Total Pts.		434	144	0
Mean		6.88	4.8	0

The results of the above set of tables show about the same thing as the set for the Cleveland School; namely, with each new test more pupils make a higher score than they did the first time, and the I. Q. varies some, both in losses and gains, but the gains are greater than the losses.

In comparing test three with test one of the Lowell School, Table II, 81 pupils, about 80 per cent, made a gain of almost 12 points each; whereas in the Cleveland School, with the same tests, 219 pupils, about 66 per cent, made gains of only 9 points.

In test three for the Lowell School, Table II, the total points gained were 943; the total lost, 95. This makes 1038 for 98 pupils. This gives an average change of over 10.5 points. Again, this is greater than the average change of 8.11 points of the Cleveland School.

Subtracting the total lost points, 95, from the total gained points, gives 848, the total net gain. This gives an average net gain of 8.65 points. Then each child of the Lowell School gained twice as many points of intelligence as each child in the Cleveland School. This somewhat substantiates the writer's opinion that the I. Q. is less constant where life and school are more variable.

Though the average change in I. Q., 8.11 in one school and 10.5 in the other, is not nearly so great as in the Stoddard study or the Woolley study, the fact that one school made twice the net gain as the other would suggest to the writer that the I. Q. might vary more in schools where school and life were more variable, where any of the factors men-

tioned in the chapter about environment, or where any of the factors mentioned in the chapter about intelligence tests not being accurate measurements, might have a chance to yield their influence.

After all, it must be remembered that the above average changes of 8.11 and 10.5 are for groups, not for individuals. Speaking of individual variation, Terman[9] has this to say:

> An I. Q. of 85, for example, means no more or less than that the child tested later will probably be found between 80 and 90. It does not mean that he may not later test as high as 100 or as low as 70, although the chances are roughly 22 to 1 against his doing so.

Though Terman does not think the number of extreme variations is very great, the writer concludes from the study of the above schools that the chances for extreme variation and an increase in total variation might be greater in schools where the factors of school are more variable than they were in the Tulsa schools studied or in the schools in which Terman did his retesting. In other words, it is the writer's opinion, based upon the findings of the two somewhat different schools in Tulsa, that if three different intelligence tests were given in one of Oklahoma's average rural high schools there would be a chance for greater average variation or change than that found in the above study.

Miss Holmes, director of tests in the Tulsa schools, said that a few years ago there were many more variations in I. Q. scores than now. She said that ever since about

---

9. L. M. Terman, Op. Cit., p. 300

five years ago that she and her department had been doing almost everything possible to weed out the extreme variations, as far as administering the tests was concerned, and had found that one of the chief causes was children giving their ages incorrectly. She made the statement that with doing everything they could to avoid it, that even now at least 5 per cent of the Tulsa I. Q. scores were inaccurate because of this factor. She added that birth certificates were the only solution to their problem and that help from them for the upper grades would not come for several years.

The above information was given to show that the Tulsa schools have one of the best testing departments in Oklahoma. Yet, 5 per cent of their I. Q.'s are inaccurate because of incorrect ages, not to say anything about the dozens of other factors that might affect the I. Q. Besides this, in the Cleveland Junior High School, the above study showed a group change of 8.11, not to say anything of the individual changes of greater variation that would have to be in order to give an average change of 8.11 points for each of 321 children. And all of this under efficient testers and in an environment that is probably more constant than in the usual school, at least in small schools or rural schools.

If the I. Q. varies some for the group, and still more for the individual in the best of schools, what should be the conclusion about the constancy of the I. Q. for individuals, especially for individuals whose environment changes?

CHAPTER VI

MISCELLANEOUS FACTORS

Some of the things to be discussed in this chapter may not be so important, for there is very little information to be found in literature about some of them; however they are given here because of their relation to the more important factors, and for the information they may have for the use and interpretation of intelligence tests.

Immediately after the World War such a wave of intelligence testing swept over the country that some people thought that the giving of the tests was simply an educational fad.

> Furthermore a great deal of the early use of tests in the classroom was a matter of satisfying curiosity. Hundreds of thousands of dollars' worth of tests have been given to millions of children, taking thousands of hours of pupil and teacher time, with absolutely nothing ever coming of it.[1]

What do people think about the present giving of intelligence tests? A look at the questionnaire will help answer the question.

III.

1. Check the following that you think are detrimental influences or results of I. Q. tests in the school:

   No.

   44 (2) An educational fad

   38 (3) Given to satisfy someone's curiosity

   ___(4) Out of date

---

1. H. A. Greene and A. N. Jorgensen, The Use and Interpretations of Educational Tests, p. 333

This shows that between one-third and one-fourth, or thereabouts, still think they are an educational fad and are given to satisfy someone's curiosity. Yet, only 13 out of 154 thought they were out of date.

It the tests are not given to satisfy someone's curiosity or as an educational fad, what is done with the findings?

V.

4. After knowing the I. Q.'s of the different pupils, did you spend more, less, or about the same amount of time on your slow pupils as you did before you knew their I. Q.'s?_____

No.
17---No answer
22---Same time
0---Less time
15---More time

The above results do not accurately tell what was done after testing, for other things might have been done, but it does show that less than 23 per cent gave additional help to the ones who made low ratings. Then some good was done for the pupil with the low I. Q., but not as much as might have been.

The writer has taught in three schools where intelligence tests were given with little accomplished therefrom. In two schools the tests were given and graded but nothing was done with the results; in the third school the tests were not all graded, let alone recorded. Though the writer's experience may not be typical, and may not prove anything, it is something to be considered.

III.

1. Check the following that you think are detrimental influences or results of I. Q. tests in the school:

No.

7 (12) Wastes too much of the student's time

6 (13) Too much work on the teacher

A very small number think that the tests work the teacher too much or take up too much of the student's time.

III-1-(14) Expense of testing too great

Twenty-five checked the above, showing that several, over 16 per cent, thought that the tests were too expensive.

The following gives some more information on the above:

V.

5. Is enough benefit derived from knowing the I. Q.'s of students to compensate for the expense, the teacher's time, and the children's time spent in an I. Q. testing program? (Yes or no)_____

Yes-----28
No------14
Blank---12

Of the 54 who had given tests over half of them thought the information gained from the tests more than compensated for the teacher's time, the pupil's time, and for the expense; however, practically one-fourth did not, and one-fourth would not say.

Perhaps, when some of the above said that the expense of testing was not too great, they were thinking of giving only one test.

IV.

1. Does your school have the money and the time to give several intelligence tests, at least two or three, to every student? (Yes or no)_____

Yes-------57
No--------72
Blank-----35

In the table just before this one the one-fourth of the people who had given tests indicated that they thought the tests too expensive; in this table about one-half of all the people who answered the questionnaire, whether they had given tests or not, indicated that their schools did not have the money to give two or three intelligence tests.

This is something to think about, when many of our educators think that more than one test should be given every pupil. Tulsa gives three tests to each pupil in the junior high school, takes the average, and calls that his I. Q. Tulsa does this to make an allowance for the changing I. Q.

Peterson[2], of Peabody College for Teachers, thinks that more than one test should be given each child. Binet[3] thought that several tests should be given, even five or six.

M. R. Trabue[4], of Columbia, says:

> The tests at present available are so inadequate and crude that one who uses a single test score as the sole basis for a vital decision in the life of an American youth is guilty of most unscientific practice and possibly of a great injury to the child advised.

So, if the I. Q. is not constant and the school does not have the money to give more than one test, should it be given at all?

---

2. J. Peterson, Early Conception and Tests of Intelligence, p. 299

3. J. Peterson, Loc. Cit.

4. M. R. Trabue, Op. Cit., p. 177

If half of the schools could not buy several I. Q. tests, how many can buy both intelligence tests and achievement tests? The table below helps to answer the question.

IV.

2. Does your school have the money and time to buy and give both intelligence tests and achievement tests? (Yes or no)_____

Yes-------48
No--------67
Blank-----39

About one-third say "yes". To the two-thirds, who do not have the money and time to buy and give both achievement and intelligence tests, the writer would suggest from the following discussion that perhaps only achievement tests should be bought. Of course, the intelligence tests are better for some things than the achievement tests, but the achievement test may displace the intelligence test in many instances. This statement will be substantiated in the following discussion.

In the first place Gilliland and Jordan[5] think that classroom teachers may give achievement tests but only specialists should give intelligence tests. Holzinger[6] says intelligence tests add little to the information furnished by achievement tests. Kelley[7] says that a comprehensive achievement test measures about the same thing---90 per cent.

---

5. A. R. Gilliland and R. H. Jordan, Op. Cit., p. 39
6. W. S. Gray, Op. Cit., p. 180
7. T. L. Kelley, Op. Cit., p. 21

Not having the money to buy several intelligence tests for each child, or not having the money to buy both intelligence tests and achievement tests should not be the only reason for using achievement tests instead of intelligence tests. For prognosis in a specific subject or field Kelley[8] has showed that achievement tests are preferable to intelligence tests.

Of course achievement tests are better than intelligence tests for measuring progress in a particular school subject or in general school work.

As a basis for future classification of pupils, at least as a basis for predicting future success in school, Kelley[9] has found grades, teachers' estimates, and special tests to all be better than intelligence tests.

Even tests other than achievement may be used instead of intelligence tests. Keys[10] showed in his work on cumulative testing that an average of some 5 or 6 tests given for instructional purposes is as good as an intelligence test for sectioning and classifying pupils.

A good vocabulary test might be substituted for the intelligence test if no grades, teachers' estimates, or any other information is available to help in an immediate

---

8. T. L. Kelley, Op. Cit., p. 26

9. L. H. King, Mental and Interest Tests, p. 9

10. N. Keys, The Improvement of Measurement Through Cumulative Testing, p. 77

sectioning of a new group of pupils. Terman[11] says that a vocabulary test is a fair intelligence test. He found a .91 correlation in one of his intelligence tests between the vocabulary part and the whole test.

Again, it is probable that ratings on achievement tests do not have the possible detrimental results or after effects that intelligence tests might have. The questionnaire shows this to be true.

IV.

3. Does a low score on an achievement test stigmatize a child as much as a low score on an intelligence test? (Yes or no)_____

Yes-------19
No--------102
Blank-----33

Over 5 times as many said "no" as said "yes". Does this need further interpretation?

How many think that the results of intelligence tests stigmatize children?

III-1-(5) Stigmatizes children

Thirty-four checked the above.

V. You need not answer the following six questions unless intelligence tests have been given in your school:

1. Do the pupils in your school know what they made on their intelligence tests? (Yes or no)_____

Yes-------12
No--------33
Blank------9

Fifty-four schools are represented in the above table,

---

11. L. M. Terman, Op. Cit., p. 308

but evidently the teachers in 9 of them, for there were 9 blanks, did not know whether the children knew their I. Q. rating or not. This is probably because the tests were given and forgotten, nothing being done with the ratings. If this is true, it would not make very much difference whether the pupils in these 9 schools knew their I. Q.'s or not; but in the other schools over one-fourth knew their I. Q.'s. This proportion is too large, for if children are not to be stigmatized, they are not to find out their I. Q.'s.

> For obvious reasons, says Terman[12], the teacher should use discretion in talking about the results of tests. That the child should not be told his mental age or I. Q. has already been emphasized. The teacher will also find that it is generally unwise to discuss the test results with parents in very specific terms... ...It is best not to discuss I. Q.'s and mental ages of individual pupils too freely among acquaintances or even among colleagues. One never knows when or where a chance remark will be repeated.

The discussion to follow concerns other possible bad effects of giving intelligence tests, especially if the pupils know their I. Q.'s.

III-1

54 (6) Gives some students the "big head"

25 (7) Promotes laziness among slower group

53 (8) Promotes discord between patrons and the school

43 (9) Causes trouble among children in the same family

One-third indicated that some students get the "big head". This is a high proportion, one out of three, and is

---

12. L. M. Terman, Op. Cit., p. 299

all the more indicative when it is realized that most of the people who filled out the questionnaire were experienced teachers.

About one-sixth thought the tests promoted laziness among the slower group. The writer has heard boys in a high school in a certain college town say that they purposely made low grades on their tests so they would be put with a slower group and could have more time for sports. Whether this actually happens or not it is something to consider. Not to say anything of the parents whose children made low I. Q. ratings, in this same town the writer has heard of parents, members of the college faculty in one or two instances, complaining that because their child made a high grade on an intelligence test he was ruining his health trying to do what the teachers expected him to do in order to make an "A".

And from the questionnaire comes this information: About one-third thought that the results of intelligence tests caused trouble between the school and patrons. And it is surprising to see that 43 out of the 154 checked "Causes trouble among children in the same family".

This thesis has considered several possible detrimental influences or results of intelligence tests, and even though some of them may not be of grave importance, ten of these possible detrimental influences will be repeated, with the tabulated checks that each received in the questionnaire, as influences that, in the writer's opinion, might be avoided

by either defining the thing measured by intelligence tests as ability to do school work, or, by substituting whenever possible some other kind of test, especially the achievement test. Bearing in mind that 154 people answered the questionnaire, the tabulations repeated here ought to be fairly significant, with the exception of number (4), Out of date.

III.

1. Check the following that you think are detrimental influences or results of I. Q. tests in the school:

116 (1) Not an accurate measurement

44 (2) An educational fad

38 (3) Given to satisfy someone's curiosity

13 (4) Out of date

34 (5) Stigmatizes children

54 (6) Gives some students the "big head"

25 (7) Promotes laziness among slower group

53 (8) Promotes discord between patrons and the school

43 (9) Causes trouble among children in the same family

46 (15) Discrimination against some of the students

CHAPTER VII

THE USE OF INTELLIGENCE TESTS

Sometimes intelligence tests are given and little or nothing is done with the scores, as has been previously mentioned in this thesis, but for another citation, and from the Twenty-First Yearbook---

> Administrators reading this will in many cases be reminded of piles of unscored tests in their offices that have not received this prompt and systematic treatment.[1]

But sometimes intelligence tests are given and too much is done with the scores. Since the questionnaire has showed that of the 154 who answered it 54 represented schools where intelligence tests had been given some time or other, perhaps it might be informational to find out something about what these schools did with the I. Q. scores.

V.

2. Are your students grouped according to their M. A.'s? (Yes or no)_____

Yes--------8
No--------42
Blank-------4

Only 8, hardly 15 per cent, grouped their students according to their M. A.'s. It must be admitted that the above question is faulty; yet, even though the questionnaire did not ask if the students were classed according to I. Q.'s, A. Q.'s, or E. Q.'s, etc., it is probable the teachers would have taken "M. A." on the questionnaire as synonimous with

---

1. M. A. Miller, The Administrative Use of Intelligence Tests in the High School, p. 196

with "I. Q.", etc. At least, it can be said that of the approximately 154 schools represented, only a little over 5 per cent of them classified the students by their M. A.'s. Yet, over 35 per cent had given tests.

Even though a few classified the students according to their M. A.'s, how many kept a record of the I. Q.'s?

III.

6. Does your school have a record of the I. Q.'s of its students? (Yes or no)_____

Yes-------41
No--------84
Blank-----29

Forty-one had a record, leaving 13 that did not. This means that of all the schools represented only 26 per cent had records of the I. Q.'s. Only 35 per cent of the schools had given tests, and only 26 per cent had kept a record of the I. Q.'s. Surely this tells something about what use the present schools are making of intelligence tests.

If 8, or over 15 per cent, grouped the students according to their M. A.'s, did any of the students object to this method of chassification?

V.

3. Approximately what per cent of the students objected to being grouped according to their M. A.'s?_____

Here are the eight answers:

Answer	Per Cent
1.	0
2.	0
3.	5
4.	20
5.	25
6.	65

Answer Per Cent

7.----100
8.----"Parents prevented"

There were objections in 75 per cent of the schools where the children were classified according to their M. A.'s. In over half of the schools, more than 4, the objections were serious, and in three of them they were extremely serious. In answer eight, "Parents prevented" is very significant. This also helps to answer the topic previously discussed about whether or not intelligence tests cause trouble between the school and the patrons.

What did all of the teachers who answered the questionnaire think about the above method of classifying pupils?

III.

8. Do you believe in classifying or grouping students according to their M. A.'s? (Yes or no)_____

Yes-------64
No--------71
Blank-----19

More than 41 per cent think that the students should be classified by their M. A.'s. Yet, in actual practice less than 15 per cent so classify them.

Again, it must be remembered that students could be classified in the superintendent's office by their M. A.'s and still not be grouped by that method in their classwork, for many superintendents do not believe in homogeneous grouping by M. A.'s or by any other method.

Dr. Hill, of Harvard, said in one of his lectures at the Oklahoma A. and M. College that the only way to have

homogeneous grouping was to have one child.

Keliher[2] concluded this in his studies of homogeneous grouping:

> Homogeneous grouping, as we now have it, appears undesirable. The measurement bases requisite for such grouping presuppose its major concern with the partial, academic phases of life. Acceptance of the philosophy that education is to concern itself with the whole child means rejection of a device which selects for consideration only certain of the individual's abilities and traits. In the light of sound theory and science of education homogeneous grouping should not be employed. In the light of the evidence concerning the results proposed for grouping, it does not achieve those results. Therefore, the major conclusion is that homogeneous grouping is not desirable in our elementary schools.

Even Terman[3] says:

> Immediate and wholesale re-grading of the school on the basis of mental age as soon as the tests have been computed is not recommended.

Some people believe that even if students are grouped homogeneously it should be done by some method other than by the results of intelligence tests. For example, Hollingshead[4] found that the educational age and subject age were the best methods for classifying students for school work. Several references were made in Chapter VI to the using of achievement tests, as well as others, instead of intelligence tests for the purpose of classification.

Should the taking of intelligence tests be compulsory

---

2. A. V. Keliher, A Critical Study of Homogeneous Grouping, p. 162

3. L. M. Terman, Op. Cit., pp. 299, 300

4. A. D. Hollingshead, An Evaluation of the Use of Certain Educational and Mental Measurements for Purposes of Classification, p. 53

for all school children?

III.

7. Should all elementary and secondary schools give I. Q. tests to all the children? (Yes or no)_____

Yes-------86
No--------53
Blank-----15

Eighty-six, over half of the 154, say tests should be given, and about one-third say they should not. Though this gives no information as to how the results should be used and interpreted, it does mean that over half of the teachers thought tests should be given throughout both grades and high school.

It may be that some of the teachers who objected to the giving of tests to every child might not object to the use of them in a more restricted sense.

III.

5. If I. Q. tests are given at all, should they be more or less limited just to problem children? (Yes or no)_____

Yes-------46
No--------93
Blank-----10

But the supposition did not prove correct. Almost twice as many objected to the limiting of intelligence tests to problem children as objected to giving the tests to all children in both elementary and secondary schools.

Then, if teachers are to use the tests---

III.

2. Have teachers put too much emphasis upon the I. Q. in methods of teaching, grouping, etc? (Yes or no)

_____

Yes-------68
No--------80
Blank------6

Though more say "no" than "yes", the number that say "yes" is so great, not far from half, that it is indicative. It means that the teachers thought that intelligence tests had been misused and misinterpreted. Even some of the teachers who in table III-7 signified that the tests should be used, indicated in III-2 that too much emphasis had been put on them.

Now, comes an important question:

III.

3. Generally speaking, do you think I. Q. testing programs such as we now have should be done away with in the school? (Yes or no)_____

Yes-------31
No--------115
Blank------8

Though 31 say "yes", almost 4 times as many say "no".

What is the opinion of teachers who have had experience giving intelligence tests, about doing away with the giving of the tests?

V.

6. Regardless of all the disadvantages and drawbacks of the average I. Q. testing program, are the advantages and benefits great enough to justify such a program? (Yes or no)_____

Yes-------33
No--------10
Blank-----11

Three times as many think an intelligence testing program is worth while as those who do not. Yet, the proportion of the ones who think the tests not worth while is great

enough not to be ignored.

Stating the above question differently to all of the 154 teachers, these answers were received:

III.

4. Do you believe in some kind of an I. Q. testing program? (Yes or no)_____

Yes------139
No--------11
Blank------4

Just 11 out of 154 did not believe in any kind of an intelligence testing program, and 139, over 90 per cent, did believe in some kind of an I. Q. testing program. Then, at least one conclusion of this chapter should be that, though many of the 154 teachers objected to some of the uses of intelligence tests, almost all of them thought some kind of an I. Q. testing program should be used.

CHAPTER VIII

CONCLUSION

Before working on this thesis the writer had the conception that psychologists and leading educators thought that intelligence tests were almost perfect measurements of native endowment, but now the writer has come to the conclusion, after reading about a hundred books, articles, and pamphlets on the subject, that he, like many other teachers, not to mention people in general, had a mistaken idea to begin with, because of the popular definition of intelligence tests and what they measure. Drawing conclusions from the educational literature read for this thesis, the writer would say that very few if any of the present day authorities maintain that intelligence tests are either perfect measurements of intelligence or measurements of only native intelligence.

Chapter III gave very forcible evidence that environment affects intelligence, especially from the questionnaire. Of course the results of a questionnaire may not give conclusive evidence, but they do show what certain people believe; and since the questionnaire was filled out by teachers on questions pertaining to the teaching profession, the results should be fairly accurate. If the teachers were wrong in any of their opinions, then the results show what opposition is to be met in setting a certain wrong right.

Two different tables showed that 90 per cent or more of the teachers thought that environment affected intelligence, and quotation after quotation from leading authori-

ties was given to corroborate this evidence. Of course 90 per cent, or even a hundred per cent of the teachers, could be wrong, but it is not very likely that 90 per cent of the teachers are wrong, and such educators as Fuller, Stoddard, Wellman, Dearborn, and others.

Though the second II-1 lists environmental factors in order of their influence upon intelligence, it may give very little additional evidence that environment affects intelligence, but it is interesting to see what the teachers listed as the most influencial. However, the potency of the different factors should be interesting to school teachers, and would be something to consider in trying to raise the I. Q. of any one person or group.

Even if II-3 shows that an average of the estimates of the 154 teachers of the influence of the one environmental factor, schooling, upon intelligence is 36 per cent, the degree of influence of the different factors, or of all of the factors, is not as important in the view of the objective of this thesis as the fact that over nine-tenths of the teachers indicated that environment does affect intelligence.

Of the twelve items named in the first II-1 as possible environmental factors every one was checked not a few times but many. Not a single item received less than 50 checks out of a possible 154. This means that every one of the 12 items mentioned was considered by at least one-third of the teachers as an environmental factor which affects intelligence. So, it must be concluded, since every item received

from 50 to 136 checks, that the school teachers thought that all 12 of the possible environmental factors do affect intelligence. At least, it must be acknowledged that if any one of these twelve factors exert influence on the I. Q., intelligence is affected by environment.

All of the findings of Chapter III concerning the effects of environment upon intelligence help to show that authorities and many teachers believe that intelligence tests are not accurate measurements of only native ability, to the utter disregard of ability acquired through environment.

For one thing, this means that authorities and a great many teachers say that scores on an intelligence test should not be interpreted as indicative of pure native ability; otherwise they would say the tests are misused.

In Chapter IV some evidence was given to support the writer's opinion that the reason why such an overwhelming majority checked intelligence tests as not accurate measurements was because intelligence tests were popularly defined as measurements of only native intelligence. Thus, it was concluded that if intelligence were differently defined, perhaps not so many people would say that the tests are not accurate measurements. In other words, if it could be said that intelligence tests just simply measure intelligence, regardless of how much is the result of environment and how much is the result of something else, more people would probably think of them as being somewhat more accurate.

Since evidence from literature was given to show that

educators could not agree upon what it is that intelligence tests measure, the writer offered a practical definition---ability to do general school work. If such a definition were accepted, the old argument of the effects of environment upon intelligence might be avoided, and many of the other accusations made about the detrimental effects of intelligence tests; and then more people would accept intelligence tests as accurate measurements---measurements of ability to do general school work. Whether or not this definition is accepted for general use of the tests, it would be a sane one to use in the interpretation of test scores.

By the popular definition of intelligence and the tests, evidence taken mostly from literature and teachers' opinions was given to show that the so-called general tests are too verbal, that they do not measure power intelligence, that they do not measure the inventive mind, and that they do not give mechanically minded people and socially minded people a fair chance.

Several quotations were given to prove that the tests are not accurate measurements of future success, that I. Q. scores are poor criteria for predicting the kind and amount of success in school or life. Even if there is a positive correlation between the I. Q. and future success, it must be remembered that it is not perfect, and that many studies have been done showing that grades, subject tests, etc. are better for making prognoses.

It must be evident that the so-called general intelli-

gence tests do discriminate against some people, that they can not measure accurately the intelligence of foreigners who have an English handicap, or the intelligence of someone who has a physical handicap such as an eye defect. If this conclusion is wrong, many authorities are wrong; otherwise this means that tests are misused when given to foreigners, people with physical defects, etc.

After all, if different group intelligence tests very often do not give a score much higher than .75 in intercorrelations, evidently they do not measure the same thing or there would be a higher correlation.

And it should be remembered, in spite of the fact that there may not be enough evidence as yet to prove the theory, Thorndike may be right in thinking that each person has three I. Q.'s. And he may have more, who knows?

If no conclusive evidence was offered in the body of the thesis, especially in Chapters III and IV, to prove one way or the other some of the questions raised about the tests not being accurate measurements, at least enough evidence was given to show that some authorities and many teachers believe that the following things may be factors capable of making the tests not accurate measurements, and that they should be considered in order to prevent the misuse and misinterpretation of the scores:

1. Not given correctly
2. Not graded correctly
3. Not fair to socially minded people
4. Not fair to mechanically minded people

5. Too verbal
6. Not fair to people with an English handicap
7. Not fair to people with certain physical defects
8. Not fair to people who become "rattled" on the test
9. The effects of environment
10. The inventive mind
11. Power intelligence
12. Not an accurate prognosis
13. Can not measure character and personality
14. Low intercorrelations
15. What is measured not definitely defined

If teachers and people in general could forget the popular conception of the tests, and realize that they are not perfect measurements, there would probably be fewer objections to the use of intelligence tests, and fewer misuses and misinterpretations of the results. If the tests are not accurate because of any one or half a dozen of the above, would it not be better to define the thing measured as the writer suggested---ability to do general school work? If the so-called general intelligence tests measure only abstract intelligence instead of mechanical or social, if they measure environmental or acquired intelligence instead of native intelligence, if they do not measure character and personality, the driving force behind intelligence, and can not make an accurate prognosis, or do some of the other 15 things indirectly suggested above, it is the conclusion of the writer that it would be better to limit the measuring capacity of tests to the practical schoolman's definition of what they

measure---ability to do general school work. At least it would be a sane and practical way to use and interpret I. Q. ratings.

Almost all of the evidence that was presented to show that environment affects the I. Q. and that intelligence tests are not accurate measurements is evidence that the I. Q. is not constant. Almost any of the 15 things mentioned above as possible factors causing the tests not to be accurate measurements could be factors in causing the I. Q. to vary.

The work of Freeman and others can not be ignored, and the examples of Dearborn, and the separate studies of Woolley, Stoddard, and Woodman---all evidence that the I. Q. is not constant.

The statistical study in this thesis of the three I. Q. scores for each of the 321 pupils in the Cleveland School and for each of the 98 pupils in the Lowell School shows that the I. Q. is not constant. For every child in the Cleveland School there is an average change of 8.11 points between the first and third test. For every child in the Lowell School there is an average change of 10.5 points between the first and third test. This change is high, but not as high as the writer had expected; however, it must be remembered that this change in I. Q. is the group change, the average change, and that many of the individuals had to vary considerably in order to make a group change of 8.11 and of 10.5.

The net gain for the Cleveland School is 4.05 points, and for the Lowell School, 8.65. The fact that each child in

the Lowell School gained twice as many I. Q. points as each child in the Cleveland School is evidence that the I. Q. is less constant where school and life are more variable, where there are greater changes in environment from favorable to unfavorable situations or from unfavorable to favorable.

Considering the above mentioned 15 or more factors that might cause inaccuracies in measuring, that might cause the I. Q. to change, the writer is of the opinion that there is probably not a school in Oklahoma outside of Tulsa in which these factors might be more constant than they are in the Cleveland School; yet on the third test each child made a net gain of 4.05 points over his first test. What caused the gain in intelligence? This thesis has no proof, but the writer wonders if additional schooling or adolescence might be the cause. This is something for further study.

If there is one thing this thesis has proved it is this: The I. Q. was not constant for the pupils on which the above study was done.

Though the I. Q. may have a certain degree of constancy for the group, and a lesser degree for certain individuals, it must be remembered that educators are not agreed as to what this constancy pertains to---as to what it is that the tests actually measure.

Tests should not be given without a purpose, without having in mind some use for the scores; yet the questionnaire showed that less than 28 per cent gave additional time to the ones who made low ratings. If one of the purposes of an I. Q. testing program is to find the pupils with low I. Q. and

give them more instruction or help, evidently some testing programs are not realizing this purpose.

Most of Chapter IV and parts of Chapter V give evidence that the I. Q. changes and that each child needs to be given more than one I. Q. test. But the questionnaire indicated that half of the schools represented by the teachers filling out the questionnaire did not have the money to buy two or three I. Q. tests for each child. Evidently the expense of a testing program involving more than one testing of the entire school would be a deciding factor in many places under present financial conditions.

This thesis has considered several possible detrimental influences of intelligence tests, and after considering the information gained from literature and the questionnaire, as well as from other sources, the 9 things listed below are presented as factors that might be given some consideration in order to avoid the misuse and misinterpretation of intelligence tests. These factors are not listed as conditions that any one of which or all occur in every testing program, but as possible factors likely to occur, especially as indicated by the questionnaire.

1. Not an accurate measurement
2. Gives some students the "big head"
3. Promotes discord between patrons and school
4. Discrimination against some of the students
5. An educational fad
6. Causes trouble among children in the same family
7. Given to satisfy someone's curiosity

8. Stigmatizes children

9. Promotes laziness among slower group

Although the factors just presented are arranged in the order of the number of checks that each received in the questionnaire, the thesis has little to give that would indicate the graveness of each. In other words, this thesis says that the questionnaire indicated that in some cases the results of intelligence tests give children the "big head", but it does not say in what per cent of the cases. Thus, to determine the graveness of each one of the possible detrimental influences of I. Q. testing is something for further study.

How are these possible detrimental factors to be avoided? To answer this question was not the purpose of this thesis; however, there was some discussion on this subject in one or two of the chapters, and from this it was concluded that some of them might be avoided by either defining the thing measured by the tests as <u>ability to do general school work</u>, or, by substituting whenever possible some other kind of test, especially the achievement test.

Almost half of the questionnaires indicated that too much emphasis had been put on the tests. Evidently the tests have been misused and misinterpreted, for surely many of our educators and 68 out of 154 teachers who answered the questionnaires can not be wrong. For example, even if it is all right to give intelligence tests to all school children, it has not yet been definitely proved that children should be classified for classroom work according to their M. A.'s;

thus any teacher who gives a single test and groups children for classwork according to their M. A.'s and then gives vocational advice, etc., may be guilty of misusing and misinterpreting the tests in more ways than one.

Regardless of all the disadvantages and drawbacks of the average I. Q. testing program, regardless of all the misuses and misinterpretations of intelligence tests, and even though all most half of the teachers designated that too much emphasis had been put on I. Q. tests, 90 per cent of them indicated that there should be some kind of an I. Q. testing program.

So there should be I. Q. testing programs, but less misuses and misinterpretations.

Thus, it is the hope of the writer that some of the factors brought forth here as possible misuses and misinterpretations of intelligence tests will be considered in the future as such, with the result that many of the present detrimental results may be avoided.

Above all, may it not be said of some teacher in the future that, after giving <u>one</u> intelligence test, he gave up with---"Well, what's the use---he's a moron".

BIBLIOGRAPHY

1. Allen, C. H. Effects of Intellectual Level. Bureau of Pub., Columbia U., (1930).
2. Bagley, W. C. Determinism in Education. Baltimore, Warwick, and York, Inc., (1925).
3. Barnes, J. H. A Statistical Study of Mental Ability and Achievement of Eight and Ninth Grade Pupils of Stillwater, Oklahoma Junior High School 1924-25. Thesis, Oklahoma A. and M. College, (1928).
4. Barthelmes, H. M. The Validity of Intelligence Test Elements. Bureau of Pub., Columbia U., (1931).
5. Billett, R. O. Provisions for Individual Differences, Markings, and Promotions. B. 17, Monograph 13. Dept. of Interior, Office of Edu., (1932)
6. Book, W. F. The Intelligence of High School Seniors. McMillan Co., (1922).
7. Brueckner, L.; Melby, E. Diagnostic and Remedial Teaching. Houghton Mifflin Co., (1931).
8. Burr, M. Y. A Study of Homogeneous Grouping. Bureau of Pub., Columbia U., (1931).
9. Chen, H. S. The Comparative Coachability of Certain Types of Intelligence Tests. Bureau of Pub., Columbia U., (1928).
10. Corning, H. M. After Testing---What? Scott, Foresman, and Co., (1926).
11. Dearborn, W. F. Intelligence Tests. Houghton Mifflin Co., (1924).

12. Deffenbaugh, M. S. Use of Intelligence and Achievement Tests in 215 Cities. City School Leaflet No. 20. Dept. of Interior, Office of Edu., (March, 1925).

13. Douglass, H. R. Modern Methods in High School Teaching. Houghton Mifflin Co., (1926).

14. Drake, C. A. A Study of an Interest Test and an Affectivity Test in Forecasting Freshman Success in College. Bureau of Pub., Columbia U., (1931).

15. Fuller, R. G. Fourteen Is Too Early. Nat. Child Labor Committee, (1927).

16. Gilliland, A. R.; Jordan, R. H. Edu. Measurements and the Classroom Teacher. The Century Co., (1924).

17. Gray, W. S. Tests and Measurements in Higher Edu. The Univ. of Chicago Press, (1936).

18. Greene, H. A.; Jorgensen, A. N. The Use and Interpretation of Educational Tests. Longmann, Green, and Co., (1929).

19. Henry, T. R., "The Wandering I. Q.". Jou. of the Nat. Edu. Ass., (Feb., 1938).

20. Hines, H. C. A Guide to Educational Measurements. Houghton Mifflin Co., (1924).

21. Hollingshead, A. D. An Evaluation of the Use of Certain Educational and Mental Measurements for Purposes of Classification. Bureau of Pub., Columbia U., (1928).

22. Keliher, A. V. A Critical Study of Homogeneous Grouping. Bureau of Pub., Columbia U., (1931).

23. Kelley, T. L. Interpretation of Educational Measurements. World Book Co., (1927).

24. Keys, N. The Improvement of Measurement Through Cumulative Testing. Bureau of Pub., Columbia U., (1928).

25. King, L. H. Mental and Interest Tests. Bureau of Pub., Columbia U., (1931).

26. Madden, R. The School Station of the Hard of Hearing Child. Bureau of Pub., Columbia U., (1931).

27. Marine, E. L. The Effects of Familiarity with the Examiner upon Stanford Binet Test Performance. Bureau of Pub., Columbia U., (1929).

28. Maxwell, C. R.; Kilzer, L. R. High School Administration. Doubleday, Doran, and Co., Inc., (1936).

29. McCall, M. A. How to Measure in Education. The McMillan Co., (1927).

30. McLaughlin, J. A. A Comparative Study of the Reliability and Validity of the Artificial Language Test in the American Council Psychological Examination, 1931 and 1932 Edition. Thesis, Oklahoma A. and M. College, (1933).

31. Odell, C. W. Educational Measurements in High School. The Century Co., (1930).

32. Peterson, J. Early Conceptions and Tests of Intelligence. World Book Co., (1925).

33. Rigg, M. Some Further Data on the Language Handicap. Journal of Edu. Psy., (April, 1928).

34. Starch, D. Educational Measurements. The McMillan Co., (1918).

35. Stoddard, G., "Child's I. Q. Can Be Improved". Oklahoma City Times, (June 28, 1938).

36. Terman, L. M. The Intelligence of School Children. Houghton Mifflin Co., (1919).

37. Terman, L. M. Terman Group Test of Mental Ability. Houghton Mifflin Co., (1937).

38. Terman, L. M., and others. Intelligence Tests and School Reorganization. World Book Co., (1923).

39. The Twenty-First Yearbook. Public School Pub. Co., (1923).

40. Tiegs, E. W., "Breaking Down the I. Q." Progressive Edu., (Dec., 1936).

41. Turner, M. A Study to Ascertain the Relation Between the Intelligence Quotient and the Physical-Nutritional Condition of a Group of Junior High School Pupils. Thesis, Oklahoma A. and M. College, (1929).

42. Woody, C. Standardized Tests. American Council on Edu., (1930).

43. Young, K. An Introductory Sociology. American Book Co., (1934).

Typist:

Mr. Arnold Nored

Campus Fire Station

Phone 2480