THE WECHSLER INTELLIGENCE SCALE FOR CHILDREN
AS A PREDICTOR OF SCHOOL SUCCESS

A Thesis
Presented to
The Department of Special Education
Indiana State Teachers College

In Partial Fulfillment
of the Requirements for the Degree
Master of Science

by
Kenneth N. Orr
July, 1950
The thesis of KENNETH N. ORR

Contribution of the Graduate School, Indiana State Teachers College, Number 712, under the title —

The Wechsler Intelligence Scale For Children as a Predictor of School Success.

is hereby approved as counting toward the completion of the Master's degree in the amount of 8 hours' credit.

Committee on thesis:

[Signatures]

Rutherford B. Foster, Chairman

Representative of English Department:

[Signature]

Date of Acceptance: July 1950
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>CHAPTER</th>
<th>THE PROBLEM AND DEFINITIONS OF TERMS USED</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>I.</td>
<td>Introduction . . . . . . . . . . . . . . .</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>The problem . . . . . . . . . . . . . . .</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>Statement of the problem . . . . . . .</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>Importance of the study . . . . . . .</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>Definitions of terms used . . . . . . .</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td>Age scale . . . . . . . . . . . . . . .</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td>Binet . . . . . . . . . . . . . . . . .</td>
<td>8</td>
</tr>
<tr>
<td></td>
<td>Coefficient of correlation . . . . . .</td>
<td>8</td>
</tr>
<tr>
<td></td>
<td>Group tests . . . . . . . . . . . . . .</td>
<td>9</td>
</tr>
<tr>
<td></td>
<td>Individual tests . . . . . . . . . . .</td>
<td>9</td>
</tr>
<tr>
<td></td>
<td>Intelligence Quotient . . . . . . . .</td>
<td>9</td>
</tr>
<tr>
<td></td>
<td>Mental age . . . . . . . . . . . . . .</td>
<td>10</td>
</tr>
<tr>
<td></td>
<td>Performance tests . . . . . . . . . . .</td>
<td>11</td>
</tr>
<tr>
<td></td>
<td>Percentile rank . . . . . . . . . . . .</td>
<td>11</td>
</tr>
<tr>
<td></td>
<td>Point-scale . . . . . . . . . . . . . .</td>
<td>11</td>
</tr>
<tr>
<td></td>
<td>Raw score . . . . . . . . . . . . . . .</td>
<td>11</td>
</tr>
<tr>
<td></td>
<td>Standard deviation . . . . . . . . . .</td>
<td>12</td>
</tr>
<tr>
<td></td>
<td>Usability . . . . . . . . . . . . . . .</td>
<td>12</td>
</tr>
<tr>
<td></td>
<td>Validity . . . . . . . . . . . . . . .</td>
<td>13</td>
</tr>
<tr>
<td></td>
<td>Verbal test . . . . . . . . . . . . . .</td>
<td>13</td>
</tr>
<tr>
<td></td>
<td>Weighted scores . . . . . . . . . . .</td>
<td>13</td>
</tr>
</tbody>
</table>
II. THE GROUP STUDIED AND THE MEASUREMENTS USED

The group
The composition of the group
The levels included
SRA Primary Mental Abilities
Binet
Why it was selected
Format
Criticism of the Binet
The WISC Scale
Format
Standardization
Criticism of the WISC
The Questionnaire

III. TECHNIQUE AND RESULTS OF THE STUDY

Selection of the cases
Survey of the group
Final selection of cases
Number of cases in the final study
Administering the individual tests
Reports to the teacher
The questionnaire
<p>| Method of determining the correlation                           | 32 |
| Interpreting the correlation and the probable error            | 33 |
| Correlation of the Binet and WISC                             | 34 |
| Interscale correlation of the WISC                             | 36 |
| Questionnaire results                                         | 36 |
| Usefulness of reports                                         | 36 |
| Most value to the teacher                                     | 38 |
| Least value to the teacher                                     | 38 |
| The more accurate measurement                                 | 38 |
| Teachers' opinion of the Binet's ability to predict school success | 40 |
| Teachers' opinion of the WISC's ability to predict school success | 40 |
| Preference for reporting test results                         | 40 |
| School marks and test predictions                             | 43 |
| IV. SUMMARY AND CONCLUSIONS                                   | 46 |
| Summary                                                       | 46 |
| Conclusions                                                   | 49 |
| BIBLIOGRAPHY                                                  | 51 |
| APPENDIX A. Correspondence from the author of the WISC        | 55 |
| APPENDIX B. Correspondence from the publisher                 | 57 |</p>
<table>
<thead>
<tr>
<th>CHAPTER</th>
<th>APPENDIX C. Form letter accompanying questionnaire.</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>APPENDIX D. The Questionnaire</td>
<td>61</td>
</tr>
<tr>
<td></td>
<td></td>
<td>59.</td>
</tr>
</tbody>
</table>
LIST OF TABLES

TABLE                                                                 PAGE
I.  The Coefficients of Correlation with
    Probable errors between the Binet and
    the WISC at the Grade Levels and for
    the Total Group .................................. 35
II. Interscale correlation of the WISC ............ 37
III. Teachers' Ratings in Percentages
    of the Value of Test Results ............... 39
IV. Teachers' Percentage Rating of Test
    Results in Comparison to School Success .. 41
V.  Teachers' Rating of Preference
    for Reporting Test Results ............... 42
VI. The Comparison of School Marks and
    Test Prediction in Percentages .......... 44
CHAPTER I

THE PROBLEM AND DEFINITIONS OF TERMS USED

I. INTRODUCTION

Since Cattell in the latter years of the nineteenth century developed a test of sensory-motor ability in his search to perfect a tool capable of measuring mental capacity, and since Binet and Simon published their "clinical tool" as the results of an extensive search to detect retardation and its causes in the Paris School System, the modern psychologist and educator have been searching for the answers to the questions: What is intelligence? How can it best be measured?

Intelligence has been defined in many ways—the capacity to learn, the capacity to adjust, those tasks the mentally defective cannot do, what intelligence tests measure at a specific time under controlled conditions, etc.

Stoddard's definition is probably the most inclusive and conclusive summation:

Intelligence is the ability to undertake activities that are characterized by (1) difficulty, (2) complexity, (3) abstractness, (4) economy, (5) adaptiveness to a goal, (6) social values, (7) and the emergence of originals, and maintain such activities under conditions that demand a concentration of energy and a resistance to emotional forces.1

The key word is probably abstractness—insight, mental set, or the integration by the cerebral cortex of the stimuli received from both the organism and its environment and the association of these stimuli with past experiences to then complete the most logical, and at the same time correct, act under the conditions prevailing.

In psychometrics, one must be cognizant of the fact that the organism is in a constant state of growth and/or development, that it is also subject to change through organic or pathological degeneration, lack of motivation, and nurture; because this affective or fluid state of the organism could well be a major factor in explaining the many authoritative but seemingly different and confusing definitions of intelligence. Further, he must be aware that the organism will respond to statistical analysis which will enable him to compare the individual to his peers, and that this is generally accomplished by measurements that have evolved from the original Binet Scale—a fact which appears to relegate the conflict in defining intelligence into the realm of semantics while still enabling the psychometrist or clinician to use these tests regardless of definition.

This study was approached with the inference that this concept of intelligence neither refuted nor eliminated any important basic elements in the philosophy
II. THE PROBLEM

Statement of the problem. It was the purpose of this study (1) to investigate the validity of The Wechsler Intelligence Scale for Children as a predictor of school success, (2) to determine the relationship or correlation between the Wechsler Scale and the Revised Stanford-Binet, (3) to determine the interscale correlation of the Wechsler, and (4) to present classroom teachers' opinion of the worth of the Wechsler Scale.

Importance of the study. A newly published intelligence test has many implications in today's education of individuals who are recognized as being homogenous in species only. The sum is greater than the whole and those "parts" comprising the whole can and do exist in varying degrees. Still, psychologists have been successful in at least combining those factors essential in judging general intelligence (even though there be general lack of agreement in specificity) to the degree that "mental tests stand today as the most important single tool psychology has de-
veloped for the practical guidance of human affairs. In fact, psychology has been so successful in formulating these measurements that an examination of any Test Bureau's catalogs will reveal an apparent abundance; but, as Cronbach remarks, both industry and education are "... learning that the test which best solves one problem may be of little value in other situations." While striving to be fair to the individual being evaluated, the clinical worker must avail himself of the test which is best for the task at hand. At the same time he must remember that certain tests are superior to others for specific jobs. "There is no one test of any sort which is 'best' for all purposes." It must be remembered that "the psychic life of man is determined by his goal." 

Intelligence tests can bring out more than just a "score", how a pupil ranks or rates with his classmate. They also enable the educator to prevent frustrations and conflicts of mental health and to adjust the general cur-

---


4 Ibid., p. xi.

5 Ibid., p. 43.

riculum to the individual.

To substantiate this inference, Axline states:

The behavior of the individual at all times seems to be caused by one drive, the drive for complete self-realization. When this drive is blocked by pressure from without, the growth toward this objective does not stop, but continues with increased momentum because of the generative force of the tensions that are created by the frustration.7

Since it is common knowledge—at least in the educational world—that one of the main objectives of education is self-realization, and since—as Hollingworth concludes—"... the most important single factor in chronic failure in school work is weaknesses in general intelligence..."8, the importance of determining the validity and usefulness of a measurement of general intelligence can be readily inferred.

Tiffin, in his discussion of the value of determining the ability of school-age children and adults, states that "... it is possible to predict intelligence at one age from measurements at an early age with a high degree of accuracy..."9 He continues that this prediction of ac-

curacy"... varies between a correlation of around .90 for immediate or short-time prediction to one of around .60 to .70 for long-time predictions.10

Lunn11 in discussing the value of knowing the intelligence quotient in educational practice stresses the following: if the intelligence quotient is lower than that required for success in the school system, the pupil can be removed from the frustrations of competing in regular classes; if the intelligence quotient is exceptionally high (a very important factor, in achieving or losing the objective of self-realization which is not receiving proper emphasis in most school systems), the pupil can be given opportunities consonant with his abilities; the findings of intelligence tests can also be used in vocational guidance.

Schonell presents the problem in epitome:

Only by the analysis of individual differences can the nature and implications of backwardness [also the rapid-learner and the gifted] in school be understood and provisions be made for their diagnosis and treatment...12

---

10 Loc. cit.


12 Frederick J. Schonell, Backwardness In The Basic Subjects (London: Oliver and Boyd, 1942) p. 3.
The measurement most often used in clinical work is highly verbal in content and not capable of clinical diagnosis of differences in kind and quality of the total pattern of intelligence (Cf. post, p. 19); therefore, the Wechsler Intelligence Scale for Children with its three quotients, total intelligence, verbal intelligence, and performance intelligence, appears most promising.

To move the importance of such a study from a philosophical approach to the basic realm of reality, Ross makes the often forgotten statement, "... there is really one fundamental purpose in all measurement: namely, the better understanding of the individual pupil."14

III. DEFINITIONS OF TERMS USED

Age scale. In this type of measurement, the tests or tasks are standardized according to the per cent passing at each age level or the "average performance of persons of each age."15 After proper standardization determines the age placement of the task, the quantity that the individual must complete to receive credit at this placed level must

15 Cronbach, *op. cit.*, p. 75.
next be determined. For example, if standardization has determined that a score of five is needed to "pass" a task at year six, and if an individual has successfully accomplished a score of four, then according to an age-scale measurement he has failed and earned no credit for the four tasks in which he was successful. Stoddard lists the following among the distinctive characteristics of an age-scale: multiple-group, age or year scale; selection by relation of successes to age; varied, unrelated, ungraded tests, internally standardized and inflexible; qualitative; all-or-none judgment; assumption of appearing functions; tests weighted equally.

Binet: Throughout the remainder of this study, the term will refer to The Revised Stanford-Binet Scale of 1937, Form L.

Coefficient of correlation: "... A numerical expression of the amount and direction [positive or negative] of the relationship between the two series of measures." It is a "... statistical device whereby relationship is expressed on a quantitative scale ... and it is designated by the letter r." "One of the more important uses

16 Stoddard, op. cit., p. 141 citing A Point Scale for Measuring Mental Ability, p. 156.
17 Ross, op. cit., p. 238.
... is determining the validity of a test. 19

Group tests. In terms of measurements, this type of test indicates that it can be administered to more than one individual at the same time. A group test is invaluable in surveying large groups, and—with a minimum of psychometric theory and careful attention to the published directions for administering and scoring—most teachers may use them.

Individual tests. This type of measurement differs from group tests in that it cannot be given to more than one individual at a time. In the majority of such tests, the author's directions must be closely followed. They are considered to be "clinical tools" since their use is not recommended without an extensive background in theory, supervised training, and clinical experience.

Intelligence. (Cf. Ante, pp. 1, 2.)

Intelligence Quotient. In 1912, Stern, a German psychologist, suggested that intelligence be represented as the ratio of mental age to chronological age or "mental quotient." Terman later adopted it as the I.Q. or rate of mental growth for his Stanford revisions of the Binet. 20 Cronbach states that it "... is a convenient way of summarizing" 21

19 Ross, op. cit., p. 244.
20 Ibid., p. 131.
21 Cronbach, op. cit., p. 120.
test performance. Ross believes that the mental age may be estimated from the present chronological age and I.Q. if the original intelligence quotient was computed between the years of six and thirteen. If the individual is provided average or normal educational opportunities, the I.Q. will give a probable maximum level of academic attainment.

**Mental age.** The concept of mental age was introduced by Binet in his 1908 revision of the original scale. Even though the mental age reveals the mental maturity of the individual, it will not signify the degree of brightness or dullness. The normal or average growth in mental ability for a child who is developing normally will equal approximately one month's growth for every month the individual has lived when it is averaged over a period of time.

It is an indication of mental maturity, and it would be considered in determining "readiness for first grade entry." An age-scale will immediately yield a mental age.

---

24 Ross, *op. cit.*, p. 34.
26 Ross, *op. cit.*, p. 106.
27 *Loc. cit.*
28 Ibid., p. 61.
Performance tests. Non-verbal tests, sometimes called performance tests, require the manipulation of stimuli presented in the form of pictures to be interpreted or matched, blocks with which to build indicated designs, or some other task not involving verbal response. 29

"The common performance tests probably indicate general mental ability more with younger and inferior subjects than with superior and older individuals." 30

Percentile rank.

The percentile is a description of a pupil's position in a typical age or grade group in terms of the percentage of pupils who fall below that score. ... A percentile score of 50 would, of course, be exactly at the median. 31

Point-scale. Stoddard lists the following among the distinctive characteristics of a point-scale: 32 single graded-test scale; selection by function measured; each test graded, available for wide range of ages; externally standardized, flexible; quantitative; more-or-less judgment; assumption of developing functions. (Cf. ante, p. 8)

Raw score. The raw score itself is of no significance.

---

29 Lursell, op. cit., p. 9.
30 Cronbach, op. cit., p. 165.
31 Ross, op. cit., p. 299.
32 Stoddard, op. cit., p. 141.
until it is compared with some standard. It is merely the sum earned by an individual on a test. 33 Does the running of one hundred yards in ten seconds reveal the swiftness of a runner? The answer is "yes" to one who knows the runner's age and the average time required to traverse the distance for individuals in the runner's classifications.

Standard deviation. The letters s., s.d., or the small sigma refer to the standard deviation which is a "measure of the spread of scores." 34 It is a measure of variability; "that distance above and below the mean that in a normal distribution includes 68.26 per cent of the scores." 35 "A particularly important use of the . . . standard deviation is for comparing one group with another." 36

Usability. For the purpose of this study the term will refer to the degree to which the test can be successfully employed by the classroom teachers from the reports furnished them, based upon test results and observation of the individual. Cronbach stresses this point:

The acceptability of a test to those who will use the results is equally important. . . . So important
is user acceptability in working with teachers . . . and others that the psychologist must often use a test which would be his second or third choice.37

Validity. "There are two types of validity, . . . curricular and statistical. The former is subjective, and the latter is objective."38 Both indicate "... the degree to which the test . . . measures what it claims to [measure]. In a word, validity means truthfulness."39

Statistical validity refers to the mathematical processes for determining the degree to which the test agrees with, or correlates with, some criterion which is set up as an acceptable measure of the thing in question.40

Verbal test. Tests that are verbal or classifiable as such consist of stimuli of words or mathematical symbols with the required task involving the verbal manipulation of the stimuli.41

Weighted scores. The term indicates that the average or mean score is 10, just as an I.Q. of 100 is the average or mean score. It is a standard score.42

---

37 Ibid., p. 47.
38 Ross, op. cit., p. 244.
39 Ibid., p. 65.
40 Ibid., p. 71.
41 Mursell, op. cit., p. 9.
42 Cronbach, op. cit., p. 144.
WISC. For the remainder of this study the term will refer to the Wechsler Intelligence Scale for Children.

III. ORGANIZATION OF THE REMAINDER OF THE THESIS

The remainder of this paper will present the group studies, the group test used, the selection of cases, the number included in the study, the criterion used in validating the problem, a description and explanation of the WISC, the questionnaire used in the study in order to check usability, the technique and results of the study, summary of the study, and conclusions based upon the study.
CHAPTER II

THE GROUP STUDIED AND THE MEASUREMENTS USED

I. THE GROUP

The composition of the group. The pupils involved in this study were in attendance at the Laboratory School of Indiana State Teachers College, Terre Haute, Indiana, during the academic year of 1949-50.

The Laboratory School provides instruction for nursery school through the twelfth grade level. There are two special rooms, one for crippled children, one for the mentally retarded. Even though the school is operated by Indiana State Teachers College, it is a unit of the public school system of Terre Haute.

A sum equal to three-fourths of the actual cost to the School City for an equivalent number of pupils at corresponding grade levels in the other public schools of the city is paid to the College for operating the school.

Pupils of school age who live in the area surrounding the College are required to attend the Laboratory School.

Teachers employed in the school are part of the

---

1 Margaret Lindsey, coordinator and director, "A Study of the Professional Education of Teachers at Indiana State Teachers College, Terre Haute, Indiana. Progress Report, September, 1948." (mimeograph form)
College Faculty and have academic rank corresponding to other College teachers.

A recent study\(^2\) revealed that 92 per cent of the pupils attending the school lived within the school boundaries, that 87 per cent of the pupils' families resided within the school area, that 2.7 per cent of the parents were foreign born, that those pupils not living in the prescribed school area had legally acceptable reasons, and that the majority of the school population lived in an area described as both residential and industrial.

Riggle's study concluded that "The Laboratory School appears to have a normal I.Q. distribution."\(^3\)

The levels included. It was determined that grades one, four, seven, and the special room for the mentally retarded were to be used in this analysis. This included 199 pupils out of the total school population of 817.

\(^2\) Information gathered from a perusal of committee reports, presented in 1946, subject: "The Pupil population of Laboratory School." This investigation was conducted by a group of graduate students in Education under the supervision of Dr. Kerle Brown, Department of Education, Indiana State Teachers College.

\(^3\) Mary Jo Riggle, "An Analysis of Mental Ability of Pupils from Grade One through Grade Nine of Indiana State Teachers College, Laboratory School." (unpublished Master's thesis, Indiana State Teachers College, Terre Haute, 1945), pp. 22.
II. SRA PRIMARY MENTAL ABILITIES\textsuperscript{4,5}

This group test was selected as the measuring device to be used in the initial survey of the group.

It is available in three levels: Primary, for ages five to seven; Elementary, for ages seven to eleven; and Intermediate, for ages eleven to seventeen.

The Primary battery is a twenty-four page booklet measuring five abilities: Verbal meaning; Perception; Motor; Space; and Quantitative.

The Elementary battery is an eighteen page step-down booklet measuring five abilities: Verbal meaning; Space; Reasoning; Perception; and Number. Separate answer sheets are provided. This level, as reported by the author, correlates to the degree of .76 with the Kuhlmann-Anderson Intelligence Test.

The Intermediate battery is similar to the elementary with a corresponding increase in difficulty. The abilities measured are: Verbal meaning; Space; Reasoning;

\textsuperscript{4} Thelma G. and L. L. Thurstone, "SRA Primary Mental Abilities," (Chicago: Science Research Associates, 1948) (Manual provided for each level of the battery)

III. BINET

Why it was selected. This individual Scale of general intelligence was selected as the validating instrument because, as Cronbach remarks, "There are two basic approaches to validity: logical analysis and empirical analysis." Logical analysis refers to the process of determining the precise "what" the scale measures (This study was not concerned with the precise "what", but it was concerned with one portion of the "what", Of ante p. 3). Cronbach continues, "In empirical analysis, one attempts to show that the test is correlated with some other variable and therefore measures the same thing." If analysis shows that the measurement in question correlates with the Binet scale, a proved predictor of school success, it can be assumed or concluded that it—at least in part—measures the same thing.

7 Loc. cit.
8 Ibid., p. 56.
9 Ibid., p. 114.
Heck\textsuperscript{11} states that the Binet is the measurement most often used to check the ability to cope with academic work. Louttitt\textsuperscript{12} and others\textsuperscript{13} concur and corroborate this statement.

\textbf{Format.}\textsuperscript{14} The scale is in two forms, L and M (only Form L was used in this study). The levels range from age two through age fourteen and adult levels of average, superior I, superior II, and superior III.

Below the five year level, tests are spaced at half-year intervals. Each age level is composed of six tests (with the exception of average adult, here there are eight). At the pre-school level an alternate test is available in case one of the tests is invalid.

The content of the scale includes comprehension, absurdities, word naming, drawing designs, number memory, differences and similarities, picture interpretation, vocabulary, and definition of abstract terms.

The raw score is changed to mental age and then the

\begin{itemize}
  \item \textsuperscript{13} C. O. Ross, \textit{Measurement In Today's Schools} (New York: Prentice-Hall Inc., 1947) p. 34.
\end{itemize}
rate of mental growth is indicated by an Intelligence Quotient. The mean I.Q. for the scale is approximately one hundred with a standard deviation of sixteen I.Q. points.

Criticism of the Binet. While many of the authorities contend that the Binet scale is an excellent measurement, there are many who criticise it severely. Inasmuch as it appeared to be the outstanding clinical tool used in predicting school success, it was deemed essential that some of these criticisms be presented.

Louttitt remarks that "perhaps the chief criticism of tests of the Binet type [age-scale] is that they place a decided premium on facility in the use of language."15

Cronbach16, in summarizing criticism of the Binet Scale, contends that the age-scale is inefficient because the tests are presented at a specific age level. As one progresses to a higher age level he will find that similar tests must be repeated. Further, the pupil may well experience success in certain tasks—for example vocabulary—while not reaching or attaining the required level of success to "pass" the task and still fail them as implicitly as he who fails to successfully accomplish any portion of that task. This is not the modus operandi in measurements classified as point-

16 Cronbach, op. cit., p. 134.
scales.

Munn states that "below [year six] . . . and above twelve, the Stanford-Binet Test is not so good a differentiating instrument." In discussing the problem of the constancy of the Intelligence Quotient, he proposes that perhaps "fluctuations of I.Q. above and below these levels may be attributed to this fact rather than actual changes."18

IV. THE WISC SCALE

The WISC Scale19 is an individual point-scale of intelligence designed and standardized for children age five through fifteen. It was compiled by a recognized authority in the field of measurements for mental ability and personality deviations of adolescents and adults. At present the scale is not intended as an indicator of personality tendencies, but a more complete report to be published at a later date will probably note the significance of clinical patterns as revealed by test item responses.

The concept of mental age as an indicator of basic


18 Loc. cit.

intelligence is refuted and dropped. Wechsler contends that the assumption that the mental age prescribes an absolute level of conduct meaning the same for all chronological ages that rate at that level is false, that the individual should be compared with his peers, and that the stress should be placed on the deviation from the mean of a specific age group.

This is accomplished by the "Deviation I.Q.", rather than by ranking the individual according to an I.Q. statistically determined for a composite age group.

By using this method of classification, the constancy of the I.Q. is maintained (states Wechsler) because the "person tested is assigned an I.Q. which, at his age, represents his relative intelligence rating." The mean total score for each level is set to equal an I.Q. of one hundred with a set standard deviation equal to fifteen I.Q. points. An I.Q. of one thirty-five or above will fall in the upper one per cent, sixty-five and below in the lowest one per cent, and an I.Q. of ninety to one hundred ten will include the middle fifty per cent.

Format. The WISC is composed of twelve tests subdivided into two scales: Verbal and Performance. The Full-Scale Intelligence Quotient is based on ten of the twelve tests with each sub-scale having an alternate test.

20 Ibid., p. 4.
The Verbal scale includes:

**General information.** The maximum raw-score is thirty with each question scored one or zero. The questions include:

1. How many ears have you?
10. How many things make a dozen?
20. Where is Chili?
30. What is a lien?

**General comprehension.** The maximum raw-score is twenty-eight. Each question is scored two, one, or zero. The problems include:

1. What is the thing to do when you cut your finger?
7. Why are criminals locked up?
10. Why is it generally better to give money to an organized charity than to a street beggar?

**Arithmetic.** The maximum raw-score is sixteen with each problem scored as one or zero. There is a time limit on each problem. The items include:

1. Place nine blocks in a row before the subject and say, "Count these blocks with your finger."
7. A boy had twelve newspapers and sold 5. How many did he have left? (This problem is presented orally to the subject. Calculation is accomplished without pencil and paper.)
13. Thirty-six is two-thirds of what number?
15. If a taxi charges twenty cents for the first quarter mile and five cents for each quarter mile thereafter, what will the fare for a two mile trip be? (this item—as is number 14 and 16—is read aloud by the subject and then worked mentally.)

**Similarities.** The maximum raw-score is twenty-eight. This section is divided into two parts: four analogies and
twelve similarities. Subjects under eight and those older
who are suspected of low mental ability are started with
analogies. Scoring of analogies is either one or zero.
Included are items:

1. Lemons are sour; sugar is _____.
5. In what way are a plum and a peach alike?
10. In what way are a pound and a yard alike?
16. In what way are the numbers forty-nine and
one hundred twenty-one alike?

Vocabulary. The maximum raw-score is eighty with
forty words scored two, one, or zero. The first five words
are scored two or zero. Subjects above eight years of age
are started with word number ten. If a score of two is not
achieved on words ten through fourteen, the subject is re-
turned to word nine and continues backwards until five con-
secutive two-point definitions are presented. The words to
be defined include:

1. Bicycle.
10. Diamond.
20. Microscope.
30. Affliction.
40. Traduce.

Digit span. This test was not used in the study since
it is an alternate test. The score is the sum of the highest
number of digits reproduced forward and backward without error.

Picture completion. Each picture is exposed a maxi-
mum of fifteen seconds. Response to each picture is scored
one or zero with the test being discontinued after four consecutive misses. The maximum raw-score is twenty. The pictures and their missing parts include:

1. Comb (Part of teeth missing.)
5. Cat (Whiskers missing on one side.)
10. Coat (Buttonholes missing.)
15. Profile (Eyebrow missing.)
20. House (Shadow missing.)

**Picture arrangement.** Twelve sets of pictures are presented to the subject in a random order with the instruction to arrange them to make a sensible story. The pictures start with a simple arrangement of a jig-saw puzzle. At year eight the pictures must be arranged to tell the most sensible story. All are timed. The last seven pictures are scored for correctness and time, a bonus being preferred for speed of manipulation. The maximum raw-score is fifty-seven. The arrangements include:

A. Dog (Three pictures.)
1. Fire (Four disarranged scenes.)
5. Sleeper (Five scenes.)
7. Rain (Six scenes.)

**Block design.** Ten designs using multi-colored blocks, nine in quantity, with all blocks identical. The tests at early levels require the subject to reproduce a pattern after it has been presented by the examiner. These simple designs require the use of four blocks. With design five the subject is presented with all nine blocks. The scoring involves time bonuses and accuracy. The maximum raw-score is fifty-five.

**Object assembly.** Four dissected pictures are pre-
sented to all subjects. The pictures used are a manikin, horse, face, and auto. Perfect performances are rewarded with a time bonus. The maximum raw-score is thirty-four.

Coding. For subjects under eight years the key involves such symbols as a star, ball, and a triangle with an identifying mark within their boundaries. The commission is to place these same marks in the test items. At this level the maximum raw-score is thirty-four.

For subjects eight years or older the coding involves the numbers one to nine inscribed in a large box. In the key the box is divided with a number above and the identifying symbol in the lower box. In the test boxes, the numbers are present but the lower box is blank; this blank must be filled with the correct symbol. A maximum time of two minutes is required for this test. The maximum raw-score for older subjects is ninety-three.

Mazes. This alternate test was not used in the study. Five mazes are presented to the subject. The raw-score is based on time required, and level of performance.

Standardization. The total test population, consisting of 2200 in all, comprised eleven age groups of one hundred white boys and one hundred white girls each. The children were within one and one-half months of their mid-year age. The total group ranked from age five through fifteen.

Using information provided by the United States Cen-
sus Bureau's survey of 1940, the country was divided into four geographic areas. The percentage of cases included from each area corresponded very closely to the percentage of total United States population in the area. The Urban-Rural ratio in each area was closely maintained in the sample cases, as was the ratio of parental occupations. Fifty-five institutionalized feeble-minded children were included.

**Criticism of the WISC.** Since this test is a very recent production, and since it was considered that published information criticizing it would be meager, it was decided to request bibliographical data from both the author and the publisher of the WISC. This revealed a dearth of information.

Next, it was concluded that the *Psychological Abstracts* would be the logical periodical to search for published information regarding the scale.

Perusal of the *Psychological Abstracts* for the issues of June 1949, through June 1950, revealed one reference; the WISC manual of directions.

Criticism of the WISC will depend upon future research, such as this study.

---

See appendixes A, B for copies of letters from these sources.
V. THE QUESTIONNAIRE

This method of determining the usability of the WISC was accomplished in cooperation with the Special Education Clinic, Indiana State Teachers College.

A form letter, which explained the questionnaire's objectives to the classroom teacher, accompanied the request for information on each pupil.

The questionnaire consisted of nine questions regarding the usefulness, value and accuracy of the predictions, teachers' preference for reporting intelligence test results, and a check on the actual school achievement of the pupil.

For the complete questionnaire, see appendixes C and D.
CHAPTER III:

TECHNIQUE AND RESULTS OF THE STUDY

I. SELECTION OF THE CASES

Survey of the group. All pupils in grades one, four, seven, and the special room for the mentally retarded were given the SRA Primary Mental Abilities group test. This group totaled 199 pupils. The tests were administered in the regular class rooms by the investigator and another graduate student who was conducting a study of this group test (cf. ante p. 17). The regular classroom personnel and advanced students in Special Education in a class on Mental Measurements were used as monitors.

Pupils who entered these grades after the survey test was completed were not included in the study.

Profile sheets, which included the Total I.Q. and Percentile rank, were prepared for each pupil. This information was then sent to the classroom or homeroom teacher.

Final selection of cases. Inasmuch as this study was conducted by a student of Special Education, and inasmuch as the majority of problems confronting such a clinic involve mental ability—or the lack of it, it was determined that the selection of cases was to be based on the pupil's variability from the reported mean for this group test, which is an I.Q. of 100 with a standard deviation of
sixteen I.Q. points: The final selection included those pupils whose SRA Primary Mental Abilities' Total I.Q. ranked ± one and one fourth standard deviations from the mean. In intelligence quotients this was interpreted to include those pupils with an I.Q. of 120 or above and those pupils with an I.Q. of 80 or below.

This general rule was applied in all classes except the special room for the mentally retarded. Since the WISC was standardized on a group which included fifty-five feeble-minded subjects, it was decided that all pupils in this room whose age fell in the range of five through fifteen were to be included in the study. This exception to the general rule excluded three pupils because of the age factor and included one pupil whose SRA Primary Mental Abilities' Total I.Q. was 87.

Number of cases in the final study. Selection of pupils by this general rule included fifteen children in grade one, fifteen in grade four, fourteen in grade seven, and eleven in the special room for mentally retarded.

From the completion of the group test and selection of cases to the time required to complete the study, normal enrolment losses eliminated one pupil in grade four, three in grade seven, and one in the special room. The final number of cases totaled fifty.

Administering the individual tests. To eliminate
the practice factor as much as possible, it was decided that two psychometrists, the investigator and the graduate student who aided in the group test, would cooperate in administering the Binet and the WISC. Each worked with all fifty cases, and—as closely as it could be arranged—half of the cases were given the Binet first and the WISC last.

The maximum time limit was one hour for each clinical visit. Cases requiring more time were completed during a second clinical visit.

Since the pupil's behavior can indicate his interests, aspirations, and approach to solving problems, \(^1\), \(^2\), \(^3\) behavior was closely observed and recorded during the individual testing period. Interpretations of this behavior as it could affect school work and general test results were considered when predicting the probable academic success.

II. REPORTS TO THE TEACHER

A report of the psychological examination findings

---
was written after each individual test was completed. It was made a working rule that the child's performance on any test was not to be discussed until all paper work had been completed and forwarded to the classroom personnel concerned.

At this stage the cases were assigned a number with a letter suffix: A, denoted grade one; B, grade four; C, grade seven; and D, special room.

III. THE QUESTIONNAIRE

The usability of the test and teachers' opinions of the measurement were determined by sending a questionnaire for each case to the homeroom and classroom teachers involved, concerning the reported information furnished by test results.

The questionnaire was forwarded after all testing and paper work was completed for the grade level concerned.

IV. METHOD OF DETERMINING THE CORRELATION

The Spearman Rank-Difference Formula was used to determine a probable rating of the WISC at the different grade levels. Since this method involved only the position of the individual's ranking on the measurements used, it indicated the presence but not the degree of relationship

---

4 See appendixes C. and D for the complete questionnaire.
between the Binet and the WISC. Further, it was recognized that coefficients of correlation based on so few cases could not be considered too valid.

The Pearson Product-Moment Formula was employed in determining the relationships for the total group.

The probable error of coefficients of correlation that were converted from the Spearman Rank-Differences Method was obtained by the formula \( \text{PER} = \frac{.7063(1-r^2)}{\sqrt{N}} \). The formula \( \text{PER} = \frac{.6745(1-r^2)}{\sqrt{N}} \) was used in computing the probable error for the total group.

V. INTERPRETING THE CORRELATION AND THE PROBABLE ERROR

Cronbach remarks that "rarely does a validity coefficient rise above .70," and Garrett concludes that a marked relationship is indicated by a coefficient of from .40 to .70 while one of .70 to 1.00 denotes high agreement. He further states that unless the validity coefficient is

---

6 Ibid., p. 189.
is at least four times its probable error, it should not be considered reliable.\footnote{9}

\textit{Nursell}\footnote{10} notes that Wallen reported a correlation of .53 between the Arthur Point-scale of Performance Tests and the 1937 revision of the Binet. This he considered an "unusually high relationship between tests of the type involved."\footnote{11}

\section{VI. CORRELATION OF THE BINET AND WISC}

Table I, page 35 lists the validity coefficients and probable errors found between the Binet and the WISC Full-scale. For the total group this was \( .77 \pm .04 \). The WISC Performance-scale and the Binet validity coefficient was \( .61 \pm .06 \). A correlation of \( .38 \pm .02 \) was indicated between the WISC Verbal-scale portion of the measurement and the Binet.

At the first grade level the WISC Full-scale correlated with the Binet \( .77 \pm .08 \). The agreement between the Performance-scale and the Binet at this level was \( .62 \pm .12 \). A correlation of \( .63 \pm .09 \) was found between the Verbal-scale and the Binet for the first grade.

The validity coefficients of correlation at the

\footnote{9} \textit{Ibid.}, p. 170

\footnote{10} \textit{Nursell, op. cit.}, p. 171 citing \textit{Journal of Genetic Psychology. 69:45-55, 1946}

\footnote{11} \textit{Loc. cit.}
TABLE I

THE COEFFICIENTS OF CORRELATION WITH PROBABLE ERRORS BETWEEN THE BINET AND THE WISC AT THE GRADE LEVELS* AND FOR THE TOTAL GROUP**

<table>
<thead>
<tr>
<th></th>
<th>Verbal Scale</th>
<th>Performance Scale</th>
<th>Full Scale</th>
</tr>
</thead>
<tbody>
<tr>
<td>First grade</td>
<td>.63±.09</td>
<td>.62±.12</td>
<td>.77±.08</td>
</tr>
<tr>
<td>Fourth grade</td>
<td>.54±.11</td>
<td>.55±.13</td>
<td>.67±.10</td>
</tr>
<tr>
<td>Seventh grade</td>
<td>.68±.05</td>
<td>.68±.12</td>
<td>.79±.08</td>
</tr>
<tr>
<td>Special room</td>
<td>.81±.08</td>
<td>.49±.17</td>
<td>.71±.11</td>
</tr>
<tr>
<td>Total Group</td>
<td>.88±.02</td>
<td>.61±.06</td>
<td>.77±.04</td>
</tr>
</tbody>
</table>

*Because of the number of cases at each level (grade one, 15; grade four, 14; grade seven, 11; and special room, 10) the Spearman Rank-Differences Method was used. The p was converted to r by the use of Table XX in Statistics in Psychology and Education, Henry E. Garrett, (New York: Longmans, Green and Co., 1926) p. 192.

The Pearson Product-Moment Formula was used.
fourth grade level were: WISC Full-scale, .67±.10; Performance-scale, .55±.13; Verbal-scale, .64±.11.

The relationship between the Binet and WISC at the seventh grade level was: Full-scale, .79±.08; Performance-scale, .66±.12; Verbal-scale, .68±.05.

Agreement in the special room for the mentally retarded was indicated by a correlation of .71 and a PE of .11 for the Full-scale, a correlation of .49 and a PE of .17 for the Performance-scale, and a correlation of .81 and a PE of .06 for the Verbal-scale.

VII. INTERSCALE CORRELATION OF THE WISC

The coefficients of correlation for the subscales with the Full-scale are noted in Table II, page 37.

The agreement between the Full-scale and the Verbal-scale was a correlation of .91 and a probable error of .02. The Full-scale and the Performance-scale correlated .93±.01. A correlation of .75±.04 was found between the Verbal-scale and the Performance-scale.

VIII. QUESTIONNAIRE RESULTS

Usefulness of Reports. Ninety per cent of the questionnaires were completed for this question. The findings concerning the general value of the tests were that ninety-two per cent of those replying felt that the Binet
### TABLE II

**INTERSCALE CORRELATION OF THE WISC**

<table>
<thead>
<tr>
<th>Scale</th>
<th>Verbal Scale</th>
<th>Performance Scale</th>
</tr>
</thead>
<tbody>
<tr>
<td>Verbal-scale</td>
<td>.73±.04</td>
<td>.91±.02</td>
</tr>
<tr>
<td>Performance-scale</td>
<td>.73±.04</td>
<td>.93±.01</td>
</tr>
<tr>
<td>Full-scale</td>
<td>.91±.02</td>
<td>.93±.01</td>
</tr>
</tbody>
</table>
report was useful while one hundred per cent reported the WISC as useful.

Lost value to the teacher. All fifty questionnaires were completed for this item. Twenty-two per cent reported the Binet as contributing the most aid to the teacher. Fifty per cent felt that the WISC was the most helpful. Twenty-two per cent reported the information of both measurements to be of equal value. This question and others of the questionnaire did not total one hundred per cent, because one other test was evaluated with the same questionnaire.

Least value to the teacher. Ninety per cent response was received on this item. Nine per cent reported the Binet as the least value while sixteen per cent classified the WISC as such. The majority of the responses listed the test which was not reported in this study as the least valuable. Approximately two per cent felt that the item was not answerable.

Table III, page 39, shows the teachers' rating of the general and comparative value of the measurements.

The more accurate measurement. This question, an attempt to determine which test agreed more closely with the teachers' opinion of the pupil's ability, was completed in forty-nine of the fifty questionnaires. Response indicated that twenty per cent favored the Binet, while forty-nine per cent believed that the WISC gave a more accurate picture
TABLE III

TEACHERS' RATINGS IN PERCENTAGES OF THE VALUE OF TEST RESULTS*

<table>
<thead>
<tr>
<th></th>
<th>General Value</th>
<th>Comparative Value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Not Useful</td>
<td>Useful</td>
</tr>
<tr>
<td>Binet</td>
<td>92</td>
<td>8</td>
</tr>
<tr>
<td>WISC</td>
<td>100</td>
<td>0</td>
</tr>
</tbody>
</table>

*The percentage does not total 100 per cent in all classifications because one other test was evaluated but not reported in this study.
of the pupil's ability to cope with school work, and twenty-two per cent of the responses reported the two tests to be of equal accuracy.

Teachers' opinion of the Binet's ability to predict school success. Ninety-six per cent response was received on this item with four per cent stating that the Binet predicted school success too low, eighty-three per cent reported it accurate, while twelve per cent stated the prediction was too high.

Teachers' opinion of the WISC's ability to predict school success. One hundred per cent response was received for this question. Six per cent felt that the WISC predicted school success too low, eighty per cent felt its predictions to be accurate, twelve per cent estimated the pupil's ability as lower than was predicted, while two per cent marked the item as questionable.

For a comparison of teachers' rating of the test results and their opinion of the pupil's ability see Table IV, page 41.

Preference for reporting test results. Since the WISC refutes the use of mental age as a method of reporting the level of mental growth, and since it is used by the Binet, the validating criterion, as a major means of reporting test results, it was deemed advisable to sample the teachers'
TABLE IV

TEACHERS' PERCENTAGE* RATING OF TEST RESULTS IN COMPARISON TO SCHOOL SUCCESS

<table>
<thead>
<tr>
<th>Rating</th>
<th>Binet</th>
<th>WISC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Too Low</td>
<td>4</td>
<td>6</td>
</tr>
<tr>
<td>Accurate</td>
<td>84</td>
<td>80</td>
</tr>
<tr>
<td>Too High</td>
<td>12</td>
<td>12</td>
</tr>
<tr>
<td>Questionable</td>
<td>0</td>
<td>2</td>
</tr>
</tbody>
</table>

*The percentage does not equal 100 per cent because one other test was evaluated but not reported in this study, and also because some responses stated that the tests were of equal value.
<table>
<thead>
<tr>
<th>Method of Reporting Results</th>
<th>Rank</th>
<th>Total Score</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>Total I.Q.</td>
<td>28</td>
<td>5</td>
</tr>
<tr>
<td>Verbal I.Q.</td>
<td>3</td>
<td>11</td>
</tr>
<tr>
<td>Mental Age</td>
<td>6</td>
<td>16</td>
</tr>
<tr>
<td>Performance I.Q.</td>
<td>0</td>
<td>5</td>
</tr>
<tr>
<td>Percentile Rank</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

*The total score was derived by multiplying the rank times the number of times the method was listed at that rank. (e.g. Total I.Q. was ranked first—28x1, second—5x2=10, third—3x3=9, fourth—1x4=4, and fifth—0x5=0. The score is (28+10+9+4+0=51)*
preference for receiving test results. Table V, page 42, tabulated the findings on this question.

The response was seventy-five per cent complete on this item with Total I.Q. ranked first, Verbal I.Q. ranked second, Mental age third, Performance I.Q. fourth, and Percentile Rank rated last in the teachers' preference for the reporting of test findings.

**School marks and test predictions.** This item was completed on all fifty questionnaires. The results are listed in Table VI, page 44.

Here the teachers were requested to compare the pupil's marks of the latest grading period with the test-predicted degree of difficulty which would probably be encountered in academic work.

Seven subjects were listed. If the achievement in two-thirds or more of the subjects corresponded with the test predictions, the results were considered accurate, accurate with higher achievement in special subjects, or accurate with specific subject disability.

School marks were lower than predicted in eight percent of the cases on the basis of Binet results and in six per cent of the cases according to WISC results.

The Binet prediction was accurate in sixty-six per cent of the cases, and the WISC achieved seventy-four per cent accuracy.
TABLE VI
THE COMPARISON OF SCHOOL MARKS AND TEST PREDICTION IN PERCENTAGES

<table>
<thead>
<tr>
<th>Achievement</th>
<th>Binet</th>
<th>WISC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lower than predicted</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>As predicted</td>
<td>66</td>
<td>74</td>
</tr>
<tr>
<td>Higher than predicted</td>
<td>10</td>
<td>6</td>
</tr>
<tr>
<td>As predicted with specific disability*</td>
<td>6</td>
<td>3</td>
</tr>
<tr>
<td>As predicted with higher achievement in special subjects*</td>
<td>10</td>
<td>6</td>
</tr>
</tbody>
</table>

*The subjects most mentioned were English, reading, and arithmetic.
Ten per cent achieved greater success than the Binet predicted. Six per cent of the cases surpassed the WISC predictions.

Special subject disability was being experienced by six per cent of the cases reported by the Binet. Eight per cent were experiencing specific subject disability, but the general level of achievement was as predicted by the WISC.

Ten per cent and six per cent of the cases were experiencing higher success in special subjects with their general level of achievement commensurate to that predicted by the Binet and WISC respectively.
CHAPTER IV

SUMMARY AND CONCLUSIONS

I. SUMMARY

The validity of the Wechsler Intelligence Scale for children as a predictor of school success was investigated by comparing its relationship to the Revised Stanford-Binet of 1937, Form L. and evaluating the usability of the reported test results by regular classroom teachers through the use of a questionnaire.

These criteria were selected because authorities seemed to agree that the Binet was the clinical tool most often used in the attempt to predict the pupil's probable academic success, and that the clinician's choice of measurements was often influenced by the usability or acceptability of the results by classroom teachers.

Even though there appeared to be an abundance of scales for the measurement of intelligence, it was pointed out that there is no one "best" test for all problems.

The first, fourth, seventh, and special room for the mentally retarded of the Laboratory School of Indiana State Teachers College were selected as the group to be utilized in selection of the cases for the study. This initial group totaled 199 pupils.

The SRA Primary Mental Abilities, a group test, was
used as the surveying measurement.

Final selection of cases included all pupils who scored a total I.Q. of one hundred-twenty or above and those whose total I.Q. was eighty or below. This general rule of acceptance was waived in selecting those pupils in the special room; the criterion here was the age level. The cases finally selected totaled fifty.

Two clinicians administered the Binet and WISC to these fifty cases. To eliminate the practice factor, it was decided that one half would receive the Binet first and one half would encounter the WISC first.

Reports of psychological findings were forwarded to each classroom teacher and homeroom teacher who had pupils involved in the study.

After all tests and paper work were completed for a specific grade level a questionnaire which checked the test results and the usability of these results was furnished the regular teacher for each pupil involved in the study.

A validity coefficient of .77 ± .04 was found between the Binet and the WISC Full-scale. The relationship between the Binet and the WISC Verbal-scale was indicated by a correlation of .88 ± .02. The agreement between the Binet and WISC Performance-scale was a correlation of .61 ± .06.

By use of the Spearman Rank-Differences Method the
converted r indicated a high relationship between the Binet and the WISC Full-scale at all levels with the exception of the fourth grade. Here marked relationship was found. High relationship was found between the Binet score and WISC Verbal-scale in the seventh grade and special room. Marked relationship was indicated between the Binet and WISC Full-scale at the first and fourth grade levels. Marked relationship was also found between the Binet and the WISC Performance-scale at all grade levels.

The WISC Full-scale and the Verbal-scale correlated .91±.02 for the total group. A correlation of .93±.01 was determined for the WISC Full-scale and the Performance-scale. The Verbal-scale and the Performance-scale correlated .73±.04.

The questionnaire results show that one hundred per cent usefulness was indicated for the WISC while ninety-two per cent usefulness was indicated for the Binet.

The majority of the responses indicated that the reports based on the WISC were of more value to the teachers than the Binet information. Still, sixteen per cent classified the WISC results as least useful. Nine per cent felt the Binet to be of least value.

The questionnaire indicated that the teachers felt that the WISC more closely agreed with their rating of the pupil's ability in forty-nine per cent of the cases. Twenty per cent favored the Binet.
The teachers rated the Binet's ability to predict school success as accurate in eighty-three per cent of the cases. Eighty per cent accuracy was listed for the WISC.

Since the WISC's author refutes and drops the concept of Mental Age as a method of reporting general level of intellectual growth, the teachers' preference was sampled on this issue. They rated Total I.Q.; first; Verbal I.Q., second; Mental age, third; Performance I.Q., fourth; and Percentile Rank, fifth.

In comparing the pupils' school marks on the last grade received with the predicted ability to cope with school work, sixty-six per cent of the cases were as predicted by the Binet and seventy-four per cent of the cases were achieving at the level predicted by the WISC.

II. CONCLUSIONS

Since consensus of opinion appeared to be that a high degree of relationship with a proved criterion of the problem under investigation indicated that the tests involved were probably measuring the same thing, and since authorities stated that a validity coefficient of .70 to 1.00 denoted a high relationship, it was concluded that the WISC can be used as a criterion of school success.

Further, inasmuch as there was found to be a high degree of agreement between the WISC Verbal-scale and the Binet,
it was inferred that this sub-scale can be used interchangeably with the Binet or in place of the Binet with a great saving in time.

In view of the statement by Lursell that a reported correlation of .53 between the Binet and the Arthur Point-scale of Performance Tests indicated an unusually high relationship between a predominantly verbal test and a performance test (OF. ante, p. 34), and since the Arthur Scale is an accepted clinical tool, it was felt that the coefficient of correlation of .61±.06 between the Binet and the WISC Performance-scale was significant, because this performance test is neither time nor space consuming.

While accepting the fact that teachers' marks and their opinion of pupils' ability to cope with school work were indeed questionable, but noting that the usability of the WISC was rated as high or higher than the Binet (the most often used clinical tool to check a pupil's ability to cope with school work) it was inferred that this validating criterion was successfully passed.

The most obvious problems raised in the research, which seemed to justify further investigation were: (1) the validation of the WISC Performance-scale with a clinically accepted Performance test of mental ability; (2) the value and use of the Mental Age as a method of reporting test results; and (3) a further subtest-analysis of the WISC.
BIBLIOGRAPHY


B. UNPUBLISHED MATERIALS


APPENDIX A.

CORRESPONDENCE FROM THE AUTHOR OF THE WISO
Dr. Rutherford B. Porter  
Director  
Special Education Clinics  
Indiana State Teachers College  
Terre Haute, Indiana  

Dear Dr. Porter:

There is as yet no bibliographical data on the Wechsler Intelligence Scale for Children and I suppose it will be some time before such data becomes available. In about a year I expect to get out a volume on the 'Mental Abilities of Children' which will serve as a handbook for the Scale.

Sincerely yours,

/s/ David Wechsler  
David Wechsler, Ph. D.  
Associate Clinical Professor  
of Medical Psychology
APPENDIX B.

CORRESPONDENCE FROM THE PUBLISHER
Dr. Rutherford B. Porter, Director
The Special Education Clinics
Indiana State Teachers College
Terre Haute, Indiana

Dear Dr. Porter:

Thank you for your letter of January 9, and especially for your kind comments with regard to the Wechsler Intelligence Scale for Children.

Because the WISC is still so new, hardly anything has appeared about it in the professional journals as yet. However, I am enclosing herewith some data which were prepared for a paper read by Dr. Seashore at the APA meeting in Denver last Fall. These same data will be the substance of an article in the Journal of Consulting Psychology in the near future. Meanwhile, Dr. Wechsler is working on a book to be devoted to the WISC, but the publication date of this is still indefinite at present.

Cordially yours,

/s/ J. H. Ricks, Jr.

James H. Ricks, Jr.
Assistant Director, Test Division

Encl.

JHR: mj
APPENDIX C.

FORM LETTER ACCOMPANYING QUESTIONNAIRE
During the present school year each of the pupils named on the enclosed questionnaires has been examined by the Special Education Clinic in regard to his or her probable academic success and achievement. The chief measures used were: The Stanford Binet Intelligence Scale; the SRA Primary Mental Abilities; and the Wechsler Scale for Children.

The individual reports on each child for these three tests have been sent to the principal's office with a duplicate copy for you. It is hoped that these reports have proved of value to you in meeting the individual needs of the pupils.

These questionnaires are submitted in order to evaluate the contribution of the findings of each test in terms of functional information to the teachers and to improve the Clinic's service to the Laboratory School. Your cooperation in answering the questions will greatly aid the Clinic in providing more useful information to you in the future.

Sincerely yours,

Rutherford B. Porter
Director

Enclosures
APPENDIX D.

THE QUESTIONNAIRE
Before marking the questionnaire you should have three reports on this child: Stanford Binet Intelligence Scale; SRA Primary Mental Abilities; Wechsler Intelligence Scale for Children.

Please check the responses which appear most fitting for _______ and return this questionnaire to the Special Education Clinic by ________.

1. Was the information in this report useful to you?
   a. Stanford Binet Report  Yes( ) No( )
   b. SRA Primary Mental Abilities Profile Sheet  Yes( ) No( )
   c. The Wechsler Scale for Children Report  Yes( ) No( )

2. For which test was the information on this report of most value to you?
   a. Stanford Binet Report  Yes( ) No( )
   b. SRA Primary Mental Abilities Profile Sheet ( )
   c. The Wechsler Scale for Children Report ( )

3. For which test was the information on this report of least value to you?
   a. Stanford Binet Report ( )
   b. SRA Primary Mental Abilities Profile Sheet ( )
   c. The Wechsler Scale for Children Report ( )

4. Which report, in your opinion gave the most accurate picture of the child's ability to learn?
   a. Stanford Binet Report ( )
   b. SRA Primary Mental Abilities Profile Sheet ( )
   c. The Wechsler Scale for Children Report ( )

5. Did the Stanford Binet Intelligence Scale predict school success too high ( ) accurately ( ) too low ( )?

6. Did the SRA Primary Mental Abilities predict school success too high ( ) accurately ( ) too low ( )?

7. Did the Wechsler Scale for Children predict school success too high ( ) accurately ( ) too low ( )?
8. Rate your preference (1, best; 2, next best, etc; 3, 4, 5) of the most beneficial means of reporting a child's standing on an intelligence test.
   a. Verbal intelligence quotient.
   b. Performance intelligence quotient.
   c. Total intelligence quotient.
   d. Mental Age.
   e. Percentile Rank

9. Please check this child's actual achievement in school work in the appropriate column.

   Very Low Below Average Average Superior Very

   English
   Arithmetic
      (numbers)
   Science
   Social Studies
   Social Science
   Reading
   Spelling