

THE STANFORD-BINET, FORM L-M, AND THE WECHSLER INTELLIGENCE
SCALE FOR CHILDREN: A COMPARATIVE STUDY UTILIZING
CULTURAL-FAMILIAL AND UNDIFFERENTIATED
MENTAL RETARDATES

APPROVED:

Earl W. Kocker

Major Professor

Jack R. Hayes

Minor Professor

Laura Lingner

Dean of the School of Education and Psychology

Robert B. Toulous

Dean of the Graduate School

THE STANFORD-BINET, FORM L-M, AND THE WECHSLER INTELLIGENCE
SCALE FOR CHILDREN: A COMPARATIVE STUDY UTILIZING
CULTURAL-FAMILIAL AND UNDIFFERENTIATED
MENTAL RETARDATES

THESIS

Presented to the Graduate Council of the
North Texas State University in Partial
Fulfillment of the Requirements

For the Degree of

MASTER OF SCIENCE

By

John S. Stone, III, B. S.

Denton, Texas

August, 1966

TABLE OF CONTENTS

	Page
LIST OF TABLES	iv
Chapter	
I. INTRODUCTION	1
Historical Background	
Development of the Tests	
Related Research	
The Use of Short Forms	
Statement of the Problem	
Hypotheses	
II. METHOD	24
Procedure	
Selection of Subjects	
Statistical Treatment	
III. RESULTS	32
Results	
Discussion	
IV. SUMMARY AND CONCLUSIONS	42
Summary and Conclusions	
Suggestions for Further Research	
BIBLIOGRAPHY	45

LIST OF TABLES

Table	Page
I. Means and the Tests of the Significance of the Difference Between Means for the Several Scales	33
II. Means, Standard Deviations, Correlation Coefficients, and the Significance of the Correlations for the Several Scales	34
III. Studies Reporting Correlations Between WISC and Stanford-Binet	39

CHAPTER I

INTRODUCTION

Historical Background

The past half century has witnessed the inception and development of psychometry, the "measurement of the duration and force of mental processes" (11, p. 598). Intelligence, being a function of one's mental processes, and its measurements are included in the field of psychometrics. Instruments for measuring intelligence have been designed for use with widely differing types of individuals. Most intelligence tests can be used with normal, above-normal, and retarded subjects. In fact, the differentiation of functional intelligence was the primary motive for designing many "IQ" tests. Many of these tests have been in use since the early stages of development of psychometric instruments. Some were revisions of earlier methods; others were more recently designed. In some instances, a test was evolved from another, possibly in order to offer an alternate technique or to fill a void in the possible choices of available instruments. Two such tests, the Stanford-Binet, Form L-M (a revised form of an earlier test) and the Wechsler Intelligence Scale for Children (a test which evolved from an earlier Wechsler test), have become the most widely used of the individual tests currently available.

Development of the Tests

The first practical and systematic approach to the problem of the differentiation of the levels of mental development was taken in 1904 in France by Binet and Simon (1). The Minister of Public Instruction named a commission to investigate possible methods of determining the educability of mentally retarded children in the French school systems. Before a child could be removed from the ordinary school and placed into a special class, he was to be given examinations to determine his ability to profit from continued education. Binet and Simon were given the task of devising an adequate method for determining childrens' intellectual development. The Development of Intelligence in Children was published by Binet and Simon in 1905 and was the first comprehensive survey of intellectual measurement. This first test consisted of thirty items arranged roughly in order of difficulty.

In 1908 the first revision of the Binet-Simon Test was published. The number of items had been increased to fifty-four and were arranged into age groups; this arrangement led to the introduction of the mental age concept which is widely used today. Binet published another revision in 1911.

Goddard (9) revised the Binet test in 1908, translated it into English and introduced it in America. Successive revisions were made by Terman (27), Terman and Merrill (28), and others (10, 12, 13, 14, 34). The 1916 Terman revision (27) was the first Stanford-Binet test and became the yardstick

by which other tests were judged for many years. The number of items had been increased to ninety and were arranged in order of difficulty by age levels. The intellectual ability of an individual was judged by comparing his performance on the scale with the standards of performance for normal children of different ages. The mental age can be defined as "the chronological age at which the average child does as well as the subject does" (5, p. 169). This scale was standardized on a sample of approximately 1,000 children and 400 adults. This revision was not only the first scale based on an "adequate" standardization of sampling practices but was also the first test to offer detailed instructions for administering and scoring each of the tests.

The 1937 revision by Terman and Merrill attempted to eliminate the major faults of the earlier Stanford-Binet. There were two forms of the 1937 revision, the Form L and the Form M. These forms were improvements, as they

1. covered a wider range of measure (by correcting the inadequate too-verbal character of the lower levels, for example);

2. were more accurately standardized (the "group consisted of 3184 native born white subjects, including approximately 100 subjects at each half-year interval from $1\frac{1}{2}$ to $5\frac{1}{2}$ years, 200 at each age from 6 to 14, and 100 at each age from 15 to 18") (28, p. 9);

3. provided a wider sampling of abilities (the number of items was increased to 192 for both forms); and

4. offered an alternate form for retesting.

Subjective judgments in scoring were minimized by the inclusion of definite principles and classified illustrations in the manual of scoring standards. In 1960, the Form L-M, another Terman-Merrill revision, was published. This single scale was made up of the best items from the L and the M forms of the 1937 scale. This L-M scale allows an alternate subtest at each level and because there is only one form, avoids duplication of items. The scale was standardized using 4498 subjects between the ages of $2\frac{1}{2}$ and 10 years.

The changes in the Form L-M are in content and in structure. The changes in content, according to the Stanford-Binet Intelligence Scale Manual (29, p. 25) include the following: (1) less satisfactory subtests were eliminated, as were duplications of retained items, (2) items that would prove more satisfactory elsewhere in the test were relocated, (3) rescoring was done where the change in difficulty was effected by a change in scoring requirements, and (4) further attempts were made to clarify and improve the directions for administering and scoring the test.

One of the several structural changes of the Form L-M adjusted the mean IQ at various age levels so that the average mental ages would more nearly equal the average

chronological ages at the various levels. Another change resulted in IQ tables with built-in adjustments for correcting atypical variability at certain age levels. At certain age levels using the 1937 Stanford-Binet, the standard deviations varied by amounts which exceeded chance expectation. It was decided that one standard deviation from the mean of a chronological age by mental age would equal sixteen IQ points, resulting in an IQ 84 for a negative deviation and an IQ 116 for a positive deviation. This deviation method gives IQ values only slightly varied from the mental age/chronological age ratio used in the 1937 IQ tables. This deviation IQ is nothing more than a standard score with the mean at 100 instead of 50 and a standard deviation of 16 instead of 10. This method, notes Pinneau, causes "a given change in score to have a constant significance as to changes in relative position regardless of age at initial test or of interval between tests" (19, p. 19).

In Measuring Intelligence (28, p. 30), which includes the manual for the Form L and the Form M, Terman stated that "the yearly gain begins to decrease after the age of thirteen and by the age of sixteen it has become approximately zero. Chronological age beyond sixteen has therefore been entirely disregarded in computing the IQ."

Later research indicated, however, that mental growth, as measured by the Stanford-Binet tests, does not cease at

age sixteen. For this reason, the IQ tables for the Form L-M were extended to include ages seventeen and eighteen.

Wechsler, the Chief Psychologist of the Bellevue Psychiatric Hospital, designed two tests for working with a wide range of patients; some of the patients were normal, some were psychotic, some were mentally retarded, and others were of miscellaneous categories of social and emotional illness. The first Wechsler test, the Form I, was published in 1939. This test was used until after World War II, when the need for an alternate form was recognized. Then, in 1946, the Form II was developed. In 1949, however, the Form II was converted into the Wechsler Intelligence Scale for Children, commonly known as the WISC.

The WISC was composed primarily of the items from the Form II of the adult scale, with easier items added to each test to permit its use with children as young as five years of age. The WISC is made up of two sections; a Verbal section consisting of six subtests; and a Performance section consisting of six subtests. If the total six subtests from either section are used, that section's scaled score must be prorated (multiplied by 5/6) because the IQ tables are based on the equivalent of five tests per section. The WISC was standardized on a sample of 2200 subjects: 100 of each sex from each age level from 5 years through 15 years.

Wechsler noted three faults or limitations to be considered when working with the "mental age" concept that had evolved from the Binet tests (31, p. 2). First, the mental age must be considered with the chronological age of the subject. A mental age of six years is not the same for a five year old child as it is for an eight year old child. Second, an IQ based on the mental age/chronological age ratio is misleading without both ages, the mental age and the chronological age of the subject, being known by the person interpreting the results. A child with a chronological age of eight years and a mental age of six years has an IQ of 75. Another child with a chronological age of twelve years and a mental age of nine years has an IQ of 75. The first child was intellectually retarded two years behind his life age whereas the second child was retarded three years behind his life age. Third, the problem of determining an adult's mental age caused Wechsler to reject Terman's earlier position that mental development ceased at age sixteen. It seemed more logical to consider adult mental age as "the age beyond which mean scores no longer increase . . .".(31, p. 3). Adult mental age is a function of the test used.

Wechsler also used the deviation IQ concept before Terman and Merrill introduced it into their Form L-M. Each age level on the WISC has its own mean and standard deviation so that, as with the 1960 Stanford-Binet, Form L-M, a person's IQ

maintains the same relative rating in any age group. The WISC was designed with a mean IQ of 100 and a standard deviation of 15.

The purpose of this study was to compare the results obtained on the Stanford-Binet, Form L-M, and the Wechsler Intelligence Scale for Children for a group of cultural-familial and undifferentiated mental retardates. Such a study should provide some evidence as to whether the two instruments adequately measure similar abilities and whether the IQ's obtained from one can be considered comparable with the IQ's obtained from the other.

Related Research

Both of these instruments, the Stanford-Binet, Form L-M, and the Wechsler Intelligence Scale for Children, are derivations of earlier tests. Some light might be thrown on the present investigation by noting the results of earlier studies that have used these and other related tests.

The 1937 Stanford-Binet, Form L, was compared with the Stanford-Binet, Form L-M, by Budoff and Purseglove (2). Their sampling consisted of seventy institutionalized mentally retarded adolescents. Three correlations were given between the Form L and the Form L-M; all IQ levels, .90; between the two forms with all subjects scoring below the IQ 50 level, .999; and between the two forms with all subjects scoring above the IQ 50 level, .92. With all IQ levels combined, the

mean of the Form L was 47.4 (S.D. 16.87) as compared with a mean of 46.46 (S.D. 16.06) for the Form L-M. The subjects with IQ's below the score of 50 had a mean of 36.1 (S.D. 5.27) on the Form L and a mean of 35.0 (S.D. 10.30) on the Form L-M. The group with IQ's above the score of 50 had a Form L of 63.8 (S.D. 7.96) and a Form L-M mean of 62.2 (S.D. 6.13). When t tests were computed to test the significance of the differences between the means, no significant differences were found. This study would indicate that the Form L-M is highly correlated with the Form L and can be used interchangeably with it when restricted to the population studied.

Several studies have been reported using the WISC and the Form L. Nale (16) found a .909 correlation between the WISC Full Scale and the Form L utilizing a group of 104 mental retardates. The mean IQ for the Form L was 55.38 (S.D. 9.85), whereas the mean IQ for the WISC-Full Scale was 57.97 (S.D. 10.15), a difference of 2.59 IQ points. These results were significant at the .001 level.

Stacy and Levin, using a group of seventy retardates, reported a correlation of .68 between the WISC Full Scale and the Stanford-Binet, Form L. The WISC Verbal Scale correlated .69 with the Form L. No information was given concerning the relationship between the WISC Performance Scale IQ and the Form L IQ or the significance of the differences between the means. The mean scores for the

various scales included 66.1 (S.D. 8.48) for the WISC Full Scale; 65.2 (S.D. 7.21) for the Form L; 66.6 (S.D. 7.0) for the WISC Verbal Scale; and 71.6 (S.D. 10.9) for the WISC Performance Scale (26).

An investigation utilizing forty retardates was carried out by Sloan and Schneider (25). The mean for the Form L was 56.3 (S.D. 4.8); the mean IQ for the WISC Full Scale was 58.3 (S.D. 9.5); the mean for the WISC Verbal Scale was 59.7 (S.D. 6.2); and the mean for the WISC Performance Scale was 64.6 (S.D. 12.7). The differences of the means of the WISC Verbal and Performance Scales were significant at the .001 level. The Form L correlated .763 with the WISC Full Scale, .751 with the WISC Verbal Scale, and .641 with the WISC Performance Scale.

Venderhost, Sloan, and Bensberg (30) studied the results obtained from thirty-eight retardates using the WISC and previous Binet tests administered within the past year. The mean Binet IQ was 59.34 (S.D. 5.07), the mean WISC Full Scale IQ was 62.18 (S.D. 7.15), the mean WISC Verbal Scale IQ was 61.74 (S.D. 7.15), and the mean WISC Performance IQ was 70.05 (S.D. 9.92). The difference between the Binet and WISC Performance Scale means was significant at the .01 level. No information was offered concerning the correlation between the two instruments.

Most of the research in the field of mental retardation concerning the 1937 Stanford-Binet was done with the Form L. The Form M was used, however, in an investigation by Sandercock and Butler (22) comparing it with the WISC and the results obtained from a group of ninety mental retarded subjects. The mean for the WISC Full Scale was 59.0 (S.D. 11.4); the mean for the WISC Verbal Scale was 62.8 (S.D. 9.7); the mean for the WISC Performance Scale was 62.6 (S.D. 12.4); and the mean for the Stanford-Binet, Form M, was 58.5 (S.D. 9.0). The differences between the Form M mean and the means of the WISC Verbal and Performance Scales were significant at the .001 level. The Form M correlated .76 with the WISC Full Scale, .80 with the WISC Verbal Scale, and .66 with the WISC Performance Scale.

Several studies have been done concerning the relationships between the Stanford-Binet, Form L-M, and the WISC using subjects of at least average intelligence. Summaries of two such studies follow. Estes, et al. (6) studied the relationships between the 1937 Stanford-Binet, the 1960 Stanford-Binet, the WISC, the Raven, and the Draw-A-Man tests. The sampling of 82 subjects including IQ's from the average group, the high average group, the superior group, and the very superior group. The average group will be the only one discussed here because of its closer proximity to the sampling included in the present study. The 1937 Stanford-Binet

scores included Form L and Form M results and their IQ's were converted into "deviation IQ's" according to the equation provided in the 1960 Stanford-Binet manual (29, p. 339). The twelve subjects that were included in the average group had a mean IQ of 100 on the 1960 Binet, a mean IQ of 102 on the 1937 Binet, and a mean IQ of 101 on the WISC Full Scale. The study notes that, "For average groups, WISC scores are comparable to both the 1937 and 1960 Stanford-Binet scores." The correlations of the tests were .82 for the 1937 Binet and the 1960 Binet, .74 for the 1960 Binet and the WISC; and .80-.90's for the 1937 Binet and the WISC. It was concluded that "at average levels of intelligence WISC scores may be used interchangeably with scores from both the S-B tests."

More recently Estes (7) correlated the scores obtained from 102 subjects who had previously tested above-average and were from homes having above-average socioeconomic status. Eighty-five subjects with IQ's ranging from "average" (actually the IQ's in this group were those considered upper-average or from IQ 100 to IQ 110), "high average," "superior," and "very superior." The difference between the means of the "average" group (Form L-M IQ=104 and WISC Full Scale IQ=105) was not significant. The Pearson product-moment correlation between the Form L-M and the WISC was .76.

Many studies and investigations, with various classifications of subjects, have utilized the WISC and the Stanford-

Binet, Form L-M. These leave little doubt that a close relationship does exist between these two instruments. Very little research, however, has been published comparing the WISC and the Form L-M utilizing a sampling of mentally retarded subjects. Poit (20), in a paper presented at the Eighty-Ninth Annual Meeting of the American Association on Mental Deficiency on June 9, 1965, notes only two published studies using mental retardates as subjects.

In one, by Burnett, the results seem to be questionable. Burnett (3) studied the relationships between the Peabody Picture Vocabulary Test (PPVT), the Wechsler scales, and the Stanford-Binet tests. All of the children at the state school were given the PPVT by a trained volunteer and the results were compared with existing results obtained from the three forms of the Stanford-Binet (Form L, Form M, and Form L-M) and the Wechsler scales (the Wechsler Intelligence Scale for Children and the Wechsler Adult Intelligence Scale). Some of the Stanford-Binet scores were as old as ten years and many of the Wechsler-Bellevue scores were two and three years old. To further complicate things, Burnett combined the Stanford-Binet scores into one group and the scores of the WISC and the WAIS into one group. He mixed "ratio IQ's" with "deviation IQ's" and combined several diagnostic groups, making no attempt to control "brain-damaged," "emotionally disturbed," "cultural-familial," or "undifferentiated" diagnoses.

The mean IQ for the Stanford-Binet results (N=238) was 61.54 (S.D. 8.59) and the range, 38-85; the mean IQ for the Wechsler Verbal scales (N=122) was 66.35 (S.D. 9.29) and the range, 44-97; the mean IQ for the Wechsler Performance scales (N=123) was 67.21 (S.D. 13.08) and the range, 40-102; and the mean IQ for the Wechsler Full Scale results (N=125) was 64.05 (S.D. 10.44) and the range, 45-93. The correlations were given as .639 between the Wechsler Verbal scales and the Binet tests (N=116); .324 between the Wechsler Performance scales and the Binet tests (N=116); and .620 between the Wechsler Full Scale scores and the Binet tests (N=117). There was no explanation why the N's making up the Wechsler Verbal, Wechsler Performance, and Wechsler Full Scale groups varied. It is difficult to understand how, when the Wechsler Full Scale scores are made up of the Wechsler Verbal and Wechsler Performance Scale scores, there can be different numbers of subjects used in the statistical manipulations. Nevertheless, Burnett notes that "All correlations significant beyond the .01 level" and that "Each pair of mean IQ's were significantly different beyond the .001 level."

The other study, by Rohrs and Haworth (21), was better designed. Much greater care was taken in controlling relevant variables. The two groups, an Organic group (N=26) and a Familial group (N=20), were cross-matched according to sex, IQ, chronological age, and length of institutionalization.

The subjects were selected by IQ level (IQ 50-70), their inclusion within the institution's academic program, the absence of visual, auditory, or physical impairment, and no exposure to any psychological test within the past year. The mean IQ's of the two groups were found to differ significantly (at the .05 level) only on the Form L-M and, therefore, "the groups were combined for further analysis." With the combined groups (N=46) the mean IQ of the Form L-M was 56.91 (S.D. 6.38) and the range, 44-71; the mean IQ of the WISC Full Scale was 52.76 (S.D. 9.70) and the range, 29-77; the mean IQ of the WISC Verbal Scale was 56.43 (S.D. 9.05) and the range, 39-84; and the mean IQ of the WISC Performance Scale was 57.54 (S.D. 11.07) and the range, 32-82. The difference between the mean IQ of the Form L-M and the mean IQ of the WISC Full Scale was significant at the .001 level. The Form L-M correlated .72 with the WISC Verbal Scale, .50 with the WISC Performance Scale, and .69 with the WISC Full Scale. All of these correlations were significant at the .01 level.

Two criticisms would seem possible concerning the Rohrs and Haworth study. First, nine subjects failed to obtain IQ's within the published WISC IQ conversion tables. These scores were extended downward according to Ogdon's extension tables. Ogdon (17) and Silverstein (24) warned that the reliability of the IQ's was attenuated by the process of extrapolation. The second criticism is that both groups, the

Familial group and the Organic group, were combined "for further analysis." Much has been written concerning the contamination of such results gained by combining heterogeneous groups such as these. The problems of "extrapolating IQ's" and combining two "heterogeneous groups" will be discussed more thoroughly in the next chapter.

The Use of Short Forms

Many institutions for the mentally retarded are understaffed. This fact has given support to recent trends for developing shorter and faster methods in determining an individual's level of intellectual functioning. The Form L-M, for example, was designed so that certain tests on each age level could be administered as a short form. The Manual notes that "each of the four tests at an age level is more heavily weighted than each of the single tests when all six tests are given" (29, p. 61). Instead of two months' credit given for each of the six tests correctly completed on the Year VI level, for example, three months' credit is given for each of four tests correctly completed. The Basal Age is then determined by the age level at which all four of the selected tests were passed.

A modified version of this Abbreviated Form was suggested by Wright (31, pp. 178-184). This method was an improvement in that the Basal Age had to include six successfully completed tests instead of Terman-Merrill's four successes, and the

Terminal Age had to include six failed tests instead of Terman-Merrill's four failures. Wright felt that it was not advisable to use any sort of modified or abbreviated scale as long as the entire Stanford-Binet could be administered. The Form L-M Manual notes that "Both methods of abbreviating the test have been found to yield somewhat lower IQ's on the average, the Terman-Merrill method slightly lower than the Weight method" (29, p. 62).

The WISC has also been the subject of attempted short forms. Carleton and Stacey found that short-form correlations ranged from .64 for a two subtest combination to .88 for a five subtest combination. The subjects were 365 children referred to a state school and considered mentally retarded or dull-normal (4).

Finley and Thompson used educable mental retardates in a study correlating the five subtest combination (Information, Picture Arrangement, Picture Completion, Coding, and Block Design) with Full Scale results. This pentad was selected by determining which combination of subtests would best predict Full Scale IQ. The multiple correlation coefficient was .896, and this was considered a valid predictor of Full Scale scores of educable mentally retarded children (8).

Two hundred and forty mental retardates were administered the WISC by Osborne and Allen (18). All possible three-subtest combinations were correlated with the WISC Full Scale.

The experiment was duplicated with a comparable group of fifty retardates. Eight three-subtest combinations were found to hold up well for both groups and were considered valid predictors of Full Scale results.

Munpower, however, found a very high two-subtest correlation (.95) with the Full Scale but when it was decided to classify each of the fifty subjects according to their short-form results, 22 per cent of the cases were misclassified when compared with their Full Scale results (15).

Another study including an intellectually heterogeneous group of children referred to a child guidance clinic found fairly poor correlations with four-subtest combinations and five-subtest combinations. The subtests used included Information, Similarities, Arithmetic, Vocabulary, Picture Arrangement, and Block Design, arranged in various groups of tests. Yalowitz and Armstrong (34) found that the highest correlation was .61 with a five-subtest combination. The writers concluded that intelligence quotients obtained from the above short-form combinations cannot be considered valid when applied to a group such as theirs.

Schwartz and Levitt, with results from 177 educable mental retardates, note that short-forms, in general, yield somewhat lower validities than earlier studies have found with normally intelligent children. Their results indicate that if a .90 correlation or better is desired, only pentads and

hexads should be used. The suggestion is given that abbreviated scales should only be given when the examiner has previous information concerning the subject's intellectual ability and that a short-form should not be used for an original intellectual evaluation (23).

These studies, both for and against the use of short or abbreviated forms, seem to imply that any initial psychological evaluation of a person's intellectual development should be undertaken with devices that have best controlled or minimized the chance for error. The WISC Full Scale and the Complete Form L-M seem to offer the greatest opportunity for valid results.

Statement of the Problem

The purpose of this study was to determine the extent of the relationship between the Stanford-Binet, Form L-M, and the Wechsler Intelligence Scale for Children for mentally retarded subjects. The results of most of the earlier studies using other subjects reported that the mean IQ scores from the Form L-M were very similar to the mean IQ scores from the three scales of the WISC. It was generally found that correlations between the Form L-M and the WISC scales were significant at least at the .05 level.

Hypotheses

The findings from the related research hypotheses suggested the following research hypotheses which were advanced concerning the expected relationships:

1. There will be no significant difference between the means of the Stanford-Binet, Form L-M, and the WISC Full Scale.
2. There will be no significant difference between the means of the Stanford-Binet, Form L-M, and the WISC Verbal Scale.
3. There will be no significant difference between the means of the Stanford-Binet, Form L-M, and the WISC Performance Scale.
4. There will be a significant, positive correlation between the Stanford-Binet, Form L-M, and the WISC Full Scale.
5. There will be a significant, positive correlation between the Stanford-Binet, Form L-M, and the WISC Verbal Scale.
6. There will be a significant, positive correlation between the Stanford-Binet, Form L-M, and the WISC Performance Scale.

CHAPTER BIBLIOGRAPHY

1. Binet, A. and Th. Simon, The Development of Intelligence in Children, Publication #11, Training School, Vineland, New Jersey, translated by E. S. Kite from various essays in L'année psychologique, Baltimore, Williams and Wilkins Co., 1916.
2. Budoff, Milton and Eleanor Mathia Purseglove, "Forms L and L-M of the Stanford-Binet Compared for an Institutionalized Adolescent Mentally Retarded Population," Journal of Clinical Psychology, XIX (1963), 214.
3. Burnett, A., "Comparison of the PPVT, Wechsler-Bellevue, and Stanford-Binet on Educable Retardates," American Journal of Mental Deficiency, LXIX (1965), 712-715.
4. Carleton, Frederick O. and Chambers L. Stacey, "Evaluation of Selected Short Forms of the WISC," Journal of Clinical Psychology, X (1954), 258-261.
5. Cronbach, Lee J., Essentials of Psychological Testing, New York, Harper & Row, 1960.
6. Estes, Betsy Worth, Mary Ellen Curtin, Robert A. DeBurger, and Charlotte Denny, "Relationships Between the 1960 Stanford-Binet, 1937 Stanford-Binet, WISC, Raven, and Draw-A-Man," Journal of Consulting Psychology, XXV (1961), 388-391.
7. _____, "Relationships Between the Otis, 1960 Stanford-Binet and WISC," Journal of Clinical Psychology, XXI (1965), 296-297.
8. Finley, Carmen J. and Jack Thompson, "An Abbreviated Wechsler Intelligence Scale for Children for Use with Educable Mentally Retarded," American Journal of Mental Deficiency, LXIII (1958), 473-480.
9. Goddard, H. H., "A Revision of the Binet Scale," Training School Bulletin, VIII (1911), 56-62.
10. Herring, John P., Herring Revision of the Binet-Simon Tests, Yonkers, World Book Co., 1924.

11. Hinsie, Leland E. and Robert Jean Campbell, Psychiatric Dictionary, New York, Oxford University Press, 1960.
12. Kuhlmann, F., "A Revision of the Binet-Simon System for Measuring the Intelligence of Children," Journal of Psycho-Asthen., Monograph Supplement, I (September, 1912), 41.
13. _____, A Handbook of Mental Tests: A Further Revision and Extension of the Binet-Scale, Baltimore, Warwick & York, 1922.
14. _____, Tests of Mental Development, Minneapolis, Educational Test Bureau, 1939.
15. Munpower, Daniel L., "The Fallacy of the Short Form," Journal of Clinical Psychology, XX (1964), 111-113.
16. Nale, Stanley, "The Children's Wechsler and the Binet on 104 Mental Defectives at the Polk State School," American Journal of Mental Deficiency, LVI (1951), 419-423.
17. Ogdon, D. P., "WISC IQ's for the Mentally Retarded," Journal of Consulting Psychology, XXIV (1960), 187-188.
18. Osborne, R. T. and J. Allen, "Validity of Short Forms of the WISC for Mental Retardates," Psychological Reports, XI (1962), 167-170.
19. Pinneau, Samuel R., Changes in Intelligence Quotient, Boston, Houghton Mifflin Company, 1961.
20. Poit, Carl H., "Wechsler and Stanford-Binet as Used with the Retardates," an unpublished paper presented at the 1965 A. A. M. D. convention in Miami, Florida, 1965.
21. Rohrs, Frederick W. and Mary R. Haworth, "The 1960 Stanford-Binet, W.I.S.C., and Goodenough Test with Mentally Retarded Children," American Journal of Mental Deficiency, LXVI (1962), 853-859.
22. Sandercock, Marian G. and A. J. Butler, "An Analysis of the Performance of Mental Defectives on the Wechsler Intelligence Scale for Children," American Journal of Mental Deficiency, LVII (1952), 100-105.

23. Schwartz, L. and E. E. Levitt, "Short-forms of the Wechsler Intelligence Scale for Children in the Educable, Non-institutionalized, Mentally Retarded," Journal of Educational Psychology, LI (1960), 187-190.
24. Silverstein, A. B., "WISC and WAIS IQ's for the Mentally Retarded," American Journal of Mental Deficiency, LXVII (1963), 617-618.
25. Sloan, William and Bernard Schneider, "A Study of the W.I.S.C. with Mental Defectives," American Journal of Mental Deficiency, LV (1951), 573-575.
26. Stacey, Chalmers L. and Janice Levin, "Correlation Analysis of Scores of Subnormal Subjects on the Stanford-Binet and W.I.S.C.," American Journal of Mental Deficiency, LV (1951), 590-597.
27. Terman, L. M., The Measurement of Intelligence, Boston, Houghton Mifflin Company, 1916.
28. _____ and M. A. Merrill, Measuring Intelligence, Boston, Houghton Mifflin Company, 1937.
29. _____, Stanford-Binet Intelligence Scale, Boston, Houghton Mifflin Company, 1960.
30. Vanderhost, L., W. Sloan, and G. J. Bensberg, Jr., "Performance of Mental Defectives on the Wechsler-Bellevue and the WISC," American Journal of Mental Deficiency, LVII (1953), 481-483.
31. Wechsler, David, Wechsler Intelligence Scale for Children Manual, New York, The Psychological Corporation, 1949.
32. Wright, Clare, "A Modified Procedure for the Abbreviated Revised Stanford-Binet Scale in Determining the Intelligence of Mental Defectives," American Journal of Mental Deficiency, XLVII (1942), 178-184.
33. Yalowitz, J. M. and R. G. Armstrong, "Validity of Short Forms of the WISC," Journal of Clinical Psychology, XI (1955), 275-277.
34. Yerkes, R. M., J. W. Bridges, and R. S. Hardwick, A Point Scale for Measuring Mental Ability, Baltimore, Warwick and York, 1915.

CHAPTER II

METHOD

Procedure

In order to investigate the hypotheses mentioned in the previous chapter, a group of fifty institutionalized mentally retarded subjects were administered both the WISC and the Stanford-Binet, Form L-M. All of the testing was done by one examiner, a staff psychologist at Whitten Village, the state school for retarded children in South Carolina.

Half of the group of fifty, twelve male and thirteen female subjects, were given the WISC on the first contact and the Form L-M on the second contact. The remaining twenty-five subjects, thirteen male and twelve female, were administered the Stanford-Binet, Form L-M, on the first contact and the WISC on the second contact. Whenever possible, the testing sessions were scheduled one week apart. For example, when the Form L-M was given on a Tuesday, the subject was recalled the next Tuesday and given the WISC. When a subject was called for a morning session, his next appointment was also scheduled in the morning. Those subjects examined during the afternoon sessions were recalled for another afternoon session.

The ten WISC subtests, five Verbal and five Performance, which are regularly scheduled, were administered to each of the subjects in this study. It was not necessary to give any of the subjects either of the alternate subtests. The standard procedures for administration of the Form L-M were also used. Most of the initial examinations were preceded by drawings of some kind; frequently a person, a tree, or a house. This was done in order to acquaint the subject with his surroundings and to aid in the establishment of adequate rapport.

Selection of Subjects

All of the subjects were residents of the institution and had diagnoses of "Mental retardation due to uncertain (or presumed psychogenic) causes with the functional reaction alone manifest." The American Association on Mental Deficiency's Manual on Terminology and Classification in Mental Retardation states that,

This category (VIII) is used for the classification of those numerous instances of mental retardation occurring in absence of any clinical or historical indication of organic disease or pathology which could reasonably account for the retarded intellectual functioning. No case is to be classified in this division except after exhaustive medical evaluation (1).

The subjects, after being included in the category (VIII) are then specifically diagnosed as either Code 81, Cultural-familial mental retardation, which requires that "there be evidence of retardation in intellectual functioning in at

least one of the parents and in one or more siblings where there are such," or Code 89, Mental retardation, other, due to uncertain cause with the functional reaction alone manifest, which requires that there is "(1) no evidence of a physical cause or structural defect, (2) no history of subnormal functioning in parents or siblings, and (3) no evidence of an associated psychogenic or psychosocial factor" (1, p. 40).

Subjects were eliminated from the study if any of the following conditions were found: (1) organicity, (2) the necessity of extrapolating WISC IQ values, (3) evidence of repeated exposure to either testing instrument, and (4) the proration of WISC IQ's. Chronological age as a variable and a determinate for selection will also be discussed.

Each patient admitted to Whitten Village generally arrives with a fairly complete medical history. All admissions have thorough social histories compiled by various social agencies. On admission, the patient is given a complete physical by one of the resident physicians. The person remains in the hospital until complete laboratory findings are analyzed. Consulting physicians are available from every field of specialized medicine. After all professions are satisfied, a medical diagnosis is made and coded according to A. A. M. D. standards for statistical purposes.

An absence of visual, auditory, or physical impairments was also considered in the selection process. No attempt,

however, was made to identify brain-damaged subjects by psychometric techniques. Haynes and Sells (4) assigned questionable validity to single variable tests, such as the Bender-Gestalt Test, the Memory-for-Designs Test, and others, when used as predictors of brain damage. Such tests "frequently identify too many false positives and false negatives, and differentiation between groups is usually gross," as well as not differentiating between the impaired efficiency of the brain and the stimulus properties of the design.

Supporting this, Beck and Lam (2), in a well-designed study found that "the WISC does not show a characteristic pattern of subtest scores for organics as a group that is representative of the individual test patterns using the method outlined by Wechsler." Wechsler's "organic test picture" noted poor Digit Span, Digit Symbol, Block Design, Arithmetic, and Similarities subtest scores. This pattern, however, was established with the WAIS, the adult scale (10, p. 217).

Extrapolation has been defined as

- 1: to infer (values of a variable in an unobserved interval) from values within an already observed interval.
- 2a: to project, extend, or expand (known data or experience) into an area not known or experienced so as to arrive at a usu. conjectural knowledge of the unknown area by inferences based on an assumed continuity . . . (9, p. 296).

Occasionally an individual would have a WISC score, either Verbal, Performance, or Full Scale, so low that the conversion tables would not include that particular score. These few subjects were dropped from the investigation. Ogdon (5) found that "extrapolated IQ values" could only be used with certain cautions or limitations. He stated that

although the empirical relationship is known to be linear, some relationships that are linear "in the middle range" of their distributions depart from linearity at the extremes.

Generally, the further the extrapolation is made from the empirical data, the greater the probability of error. In this case, as our extrapolated IQs decrease from 46, the less confidence we have in our extrapolation, the lower the expected reliability of the value, and the less it's probably psychologically significant.

Silverstein (8) devised tables of WISC Verbal and Performance IQ's that extended down to a scaled score of one. He, as did Ogdon, made certain to note that the reliability of the extrapolated IQ's is attenuated by the small numbers of items successfully completed to determine the scaled score.

The influence of repeated exposure to either of the measuring instruments was considered an important factor. All of the subjects had been examined with both instruments for this study. It was necessary, then, to eliminate from the group any subjects that had been examined with either instrument by a psychologist, including the writer, within the two years preceding the experiment. Reger (6) noted that with a group of sixty-five educable mentally retarded boys, "significant differences beyond the .01 level were

noted between the first and second administrations for the Performance quotients and for the Full Scale quotients." The mean time between the first and second administration was twelve months, with a range from two months to thirty-two months. The WISC was given a third time, again with a mean of twelve months between the second and third testing, though the range was from nine months to fifteen months. No significant differences were noted for the Verbal quotients, even between the first and second administrations. The effect of the multi-exposure on the Performance Scale evidently contaminates the results of the WISC, causing the comparison of it with the Stanford-Binet, Form L-M, to be invalidated.

Silverstein, in his study concerning the effects of prorating of WISC IQ's, reported that constant errors may be introduced by prorating the WISC IQ's of mentally retarded children (7). Wechsler noted in the WISC manual that "The general rule, then, is to give the five Verbal and the five Performance tests which are regularly scheduled unless one has clinically valid reasons for substituting an alternate" (11, p. 19).

The chronological age of the subjects was restricted by the design of the measuring instruments. The WISC conversion tables used for determining the intelligence quotients (11, p. 24) reports IQ's only as low as forty-five on the Verbal Scale, forty-four on the Performance Scale, and

forty-six on the Full Scale. In order to obtain WISC IQ's such as these without extrapolating one or more of the scales, it was necessary to select subjects with chronological ages of at least eight years. The subjects could be no older than fifteen years, eleven months, the upper limits of the WISC.

The range of chronological ages of the subjects was from eight years, eleven months to fifteen years, ten months. The mean age was 13.163 years. The male sampling (N=25) included a range from nine years, three months to fifteen years, six months with a mean age of 12.953 years. The female group (N=25) had a range of eight years, eleven months to fifteen years, ten months and a mean age of 13.373 years.

Statistical Treatment

The first three hypotheses were tested by the use of Fisher's t for correlated means. The remaining three hypotheses were tested by calculating the Pearson Product-Moment correlations between the three WISC IQ's and the Stanford-Binet, Form L-M, IQ's. The .05 level of significance was employed in both cases.

CHAPTER BIBLIOGRAPHY

1. American Association of Mental Deficiency, A Manual of Terminology and Classification in Mental Retardation 2nd ed., prepared by R. Heber, Monograph Supplement to the American Journal of Mental Deficiency, 1961.
2. Beck, Harry S. and Robert L. Lam, M. D., "The Use of the W.I.S.C. in Predicting Organisity," Journal of Clinical Psychology, XI (1955), 154-158.
3. Ferguson, George A., Statistical Analysis in Psychology and Education, New York, McGraw-Hill, 1959.
4. Haynes, Jack R. and S. B. Sells, "Assessment of Organic Brain Damage: A Critical Review," unpublished paper read at Symposium on The Assessment of Organic Brain Damage, American Psychological Association, New York, New York, September 1, 1961.
5. Ogdon, Donald P., "WISC IQs for the Mentally Retarded," Journal of Consulting Psychology, XXIV (1960), 187-188.
6. Reger, Roger, "Repeated Measurements with the WISC," Psychological Reports, XI (1962), 418.
7. Silverstein, A. B., "Effects of Proration on the W.I.S.C. IQs of Mentally Retarded Children," Psychological Reports, XII (1963), 646.
8. _____, "WISC and WAIS IQs for the Mentally Retarded," American Journal of Mental Deficiency, LXVII (1963), 617-618.
9. Webster's Seventh New Collegiate Dictionary, G. & C. Merriam Company, Publishers, Springfield, Massachusetts, 1963.
10. Wechsler, David, The Measurement and Appraisal of Adult Intelligence, Baltimore, The Williams & Wilkins Company, 1958.
11. _____, Wechsler Intelligence Scale for Children Manual, New York, The Psychological Corporation, 1949.

CHAPTER III

RESULTS

Results

As stated in Chapter One, the present research was designed to study the relationships between the Stanford-Binet, Form L-M, and the Wechsler Intelligence Scale for Children, when used with institutionalized cultural-familial and undifferentiated mental retardates. Specifically, the following theoretical hypotheses were presented: (1) That there will be no significant difference between the means of the Stanford-Binet, Form L-M, and the WISC Full Scale; (2) That there will be no significant difference between the means of the Stanford-Binet, Form L-M, and the WISC Verbal Scale; and (3) That there will be no significant difference between the means of the Stanford-Binet, Form L-M, and the WISC Performance Scale. Data relative to these hypotheses are presented in Table I.

An inspection of Table I reveals that the difference between the means of the Stanford-Binet, Form L-M, and the WISC Verbal Scale did not reach significance though it approached significance. The means of the Form L-M and the

TABLE I

MEANS AND THE TESTS OF THE SIGNIFICANCE OF THE DIFFERENCE BETWEEN MEANS FOR THE SEVERAL SCALES

M_1	M_2			$M_1 - M_2$	t	df	P
	WISC Verbal Mean	WISC Performance Mean	WISC Full Scale Mean				
Stanford-Binet Form L-M Mean	63.32	65.22	65.24	1.90	1.939	49	.10
	• • •	• • •	• • •	1.92	1.486	49	.20
	• • •	• • •	61.84	1.48	2.021	49	.05

WISC Performance Scale differed too little to be significant. The means of the Form L-M and the WISC Full Scale were significantly different ($P=.05$).

The following hypotheses previously stated concerning the significance of the correlations between the scales were

(4) That there will be a significant positive correlation between the Stanford-Binet, Form L-M, and the WISC Full Scale;

(5) That there will be a significant, positive correlation between the Stanford-Binet, Form L-M, and the WISC Verbal Scale; and (6) That there will be a significant, positive correlation between the Stanford-Binet, Form L-M, and the WISC Performance Scale. Data relative to these hypotheses are presented in Table II

An examination of Table II reveals that the Stanford-Binet, Form L-M, correlated .624 with the WISC Verbal Scale.

TABLE II

MEANS, STANDARD DEVIATIONS, CORRELATION COEFFICIENTS, AND THE SIGNIFICANCE OF THE CORRELATIONS FOR THE SEVERAL SCALES

Variable	Form L-M	Mean	Standard Deviation	df	P
Form L-M	. .	63.320	6.782
WISC Verbal Scale	.62	65.220	7.155	49	.01
WISC Performance Scale	.46	65.240	9.305	49	.01
WISC Full Scale	.68	61.840	7.190	49	.01

The Form L-M and the WISC Performance Scale correlated .457. The Form L-M and the WISC Full Scale correlated at .678. All of the above correlations were significant at the .01 level.

Discussion

The first hypothesis stated that there would be no significant difference between the means of the Form L-M and the WISC Full Scale. The findings indicate that the mean difference of the scores of these two instruments was significantly different. This makes rejection of the null hypothesis necessary. The WISC Full Scale cannot be used interchangeably with the Form L-M for this population.

The second hypothesis stated that there would be no significant difference between the means of the Form L-M and the WISC Verbal Scale. The data relating to this hypothesis reveal that the level of significance attained is .10.

Since this is a small and insignificant difference, this hypothesis was accepted.

The third hypothesis stated that there would be no significant difference between the means of the Form L-M and the WISC Performance Scale. This hypothesis was also accepted since the level of significance was only .20.

A possible reason for the difference between Stanford-Binet, Form L-M, and the WISC Full Scale has to do with how WISC IQ's are calculated. The WISC Full Scale is composed of the scaled scores of the Verbal and Performance Scales. Figuring the Full Scale score, however, is not done on a simple one-to-one basis. For example, a Verbal IQ of 60 is obtained by achieving 18 scaled score points and Performance IQ of 60 requires 21 scaled score points. The total scaled score of 39, however, yields a Full Scale IQ of only 56, 4 IQ points below the Verbal and Performance IQ scores. A similar inconsistency is found at the other end of the intelligence range. The only difference is that the Full Scale IQ is greater than the scaled score sum of the Verbal and Performance Scales.

Wechsler "predetermined that the mean IQ should be 100 and the standard deviation 15" (11, p. 15). IQ's were determined by setting the average sum of obtained scaled scores equal to 100 and the standard deviation equal to 15. This process does not greatly affect the derived IQ's until

the extremes of the normal curve are approached. Then, however, for some reason, as in the case of the mentally retarded subject, the "total" factor of intelligence is less than that person's combined verbal intelligence and non-verbal intelligence. This possibility for the rejection of the first hypothesis (that there would not be a significant difference between the means of the Form L-M and the WISC Full Scale) is based on statistical considerations in the construction of the conversion tables designed by Wechsler.

The acceptance of the second and third hypotheses suggests that the WISC Verbal and Performance Scales and the Stanford-Binet, Form L-M, measure similar qualities or traits using cultural-familial and undifferentiated retardates. At least on a group basis there are insignificant differences between means.

A finding that seems to contradict with the results of other studies was the close similarity between the mean WISC Verbal IQ score and the mean WISC Performance IQ score. Most investigators have found the Performance IQ mean score to be higher than the Verbal IQ mean score when administered to mental defectives. Alper reported that cultural-familial or undifferentiated retardates had a higher performance than verbal IQ (1). Higher WISC Performance IQ's have also been reported by Burnett (2), Guthrie and Pastovic (3), Neuman and Loos (4), Rohrs and Haworth (5), Sloan and Schneider (8),

Seashore (7), Stacey and Levin (9), and Vanderhost, Sloan, and Bensberg (10). Seashore also noted, however, that "we would be unsafe in accepting as clinically important the somewhat common generalization that the feebleminded . . . are less feebleminded on performance tests" (7).

The mean IQ difference of .02 points (the Verbal Scale being slightly higher than the Performance Scale) hardly justifies mentioning as a "difference." It would seem that, as was done in the present study, the importance of a single examiner/scorer needs to be considered a more important variable. The subjective qualities of deciding which response deserves two points' credit, one point credit, or no credit at all in the scoring of the WISC Scales and whether to give credit (a plus response) or not (a minus response) when scoring the Stanford-Binet, Form L-M, can best be controlled by using the same examiner for all subjects and having him do the scoring for all the tests.

The following hypotheses were concerned with the correlations of the various scales of the WISC with the Stanford-Binet, Form L-M. The fourth hypothesis of the study stated that the WISC Full Scale would correlate significantly and positively with the Stanford-Binet, Form L-M. This hypothesis is accepted since the correlation of .678 results in a .01 level of significance.

The fifth hypothesis stated that there would be a significant, positive correlation between the Stanford-Binet, Form L-M, and the WISC Verbal Scale. Again, the .01 level of significance was found, the correlation between the two scales being .624 and the hypothesis being accepted.

The sixth, and last, hypothesis stated that there would be a significant, positive correlation between the Stanford-Binet, Form L-M, and the WISC Performance Scale. This was accepted since the correlation of .457 was significant at the .01 level.

These correlations fall within the general trend set by other studies, including Burnett (2), Rohrs and Haworth (5), Sandercock and Butler (6), Sloan and Schneider (8), and Stacey and Levin (9). A tabular summary of the above correlations is shown in Table III.

An important factor to remember in interpreting the correlations concerns the population from which the results were obtained. Mental retardation is generally considered as the lower 3 per cent of the normal distribution. Since all of the subjects included in this study were mentally retarded, all were functioning in a very compressed range of abilities (or scores). This, of course, would affect the size of the correlations between tests.

It should be noted, in summary, that even though the correlations were significant and differences between means

TABLE III
STUDIES REPORTING CORRELATIONS BETWEEN WISC
AND STANFORD-BINET

Name	N	Correlation with Stanford-Binet		
		WISC Verbal	WISC Perform.	WISC F.S.
^a Burnett	117	.64	.32	.62
^b Rohrs and Haworth	46	.72	.60	.69
^c Sandercock and Butler	90	.80	.66	.76
^d Sloan and Schneider	40	.75	.64	.76
^d Stacey and Levin	72	.69	. .	.68
^b Stone	50	.62	.46	.68

^a Data included 1937 Binet, 1960 Binet, the WISC, WAIS, and W-B Forms 1 and 2.

^b Stanford-Binet, Form L-M.

^c Stanford-Binet, Form M.

^d Stanford-Binet, Form L.

in the two cases were not, the size of the correlations would result in considerable error if one test were used to estimate the IQ of the other and the error would be relatively greater for the Performance Scale. In other words, one should use caution in treating these IQ's as though they were equivalent for individuals.

CHAPTER BIBLIOGRAPHY

1. Alper, A. E., "An Analysis of the Wechsler Intelligence Scale for Children with Institutionalized Mental Defectives," Dissertation Abstracts, XX (1960), 4711.
2. Burnett, A., "Comparison of the PPVT, Wechsler-Bellevue, and Stanford-Binet on Educable Retardates," American Journal of Mental Deficiency, LXIX (1965), 712-715.
3. Guthrie, G. M. and J. J. Pastovic, "Some Evidence of the Validity of the WISC," Journal of Consulting Psychology, XV (1951), 385-386.
4. Newman, J. R. and F. M. Loos, "Differences Between Verbal and Performance IQs with Mentally Defective Children," Journal of Consulting Psychology, XIX (1955), 16.
5. Rohrs, F. W. and M. R. Haworth, "The 1960 Stanford-Binet, WISC and Goodenough Tests with Mentally Retarded Children," American Journal of Mental Deficiency, LXVI (1962), 853-859.
6. Sandercock, M. G. and A. J. Butler, "An Analysis of the Performance of Mental Defectives on the Wechsler Intelligence Scale for Children," American Journal of Mental Deficiency, LVII (1952), 100-105.
7. Seashore, H. G., "Differences Between Verbal and Performance IQs on the Wechsler Intelligence Scale for Children," Journal of Consulting Psychology, XV (1951), 62-67.
8. Sloan, W. and B. Schneider, "A Study of the Wechsler Intelligence Scale for Children with Mental Defectives," American Journal of Mental Deficiency, LV (1951), 573-575.
9. Stacey, Chalmers L. and Janice Levin, "Correlation Analysis of Scores of Subnormal Subjects on the Stanford-Binet and WISC," American Journal of Mental Deficiency, LV (1951), 590-597.

10. Vanderhost, L., W. Sloan, and G. F. Bensberg, "Performance of Mental Defectives on the Wechsler-Bellevue and the WISC," American Journal of Mental Deficiency, LVII (1953), 481-483.
11. Wechsler, David, Wechsler Intelligence Scale for Children Manual, New York, The Psychological Corporation, 1949.

CHAPTER IV

SUMMARY AND CONCLUSIONS

Summary and Conclusions

In this study the comparability of two widely used individual intelligence tests was examined. The results obtained from the Stanford-Binet, Form L-M, and the WISC, utilizing a group (N=50) of cultural-familial and undifferentiated retardates, were analyzed. The data suggested that the WISC Verbal Scale and the WISC Performance Scale IQ means did not differ significantly and that their correlations with the Form L-M were significant. The WISC Full Scale IQ, however, cannot be used interchangeably with the Stanford-Binet, Form L-M, since the difference between means was significant. Wechsler's statistical treatment of the data at the lower extreme of the normal distribution resulted in the WISC Full Scale IQ being several points lower than the IQ derived from averaging WISC Verbal and Performance IQ's. This may account for the significant difference between the mean IQ scores of the WISC Full Scale and the Stanford-Binet, Form L-M. The correlations of the various scales, the WISC Full Scale with the Form L-M, the WISC Verbal Scale with the Form L-M, and the WISC Performance Scale with the Form L-M, were all significant at the .01 level of confidence.

The correlations also compared favorably with the general trends set by earlier studies. The possibility that the lower correlations found with the retarded, as compared with normal subjects, were due to the compressed range of IQ scores was discussed.

Suggestions for Further Research

The importance of the data presented here is not easily assessed. Additional studies utilizing this population, possibly with a larger sample, are needed. Replication with other types of retardates would be of value. One project, for example, that should be undertaken would be similar to the present study. The purpose would be to determine the differences in functioning on the two tests by several combinations of male/female and cultural-familial/undifferentiated retardates.

Silverstein found, in a survey of ninety-six state institutions (1), that the Stanford-Binet was used a majority of the time in forty-four institutions and that the Wechsler scales were used a majority of the time by thirty-five institutions. These instruments are obviously the backbone of psychological testing in state schools for the retarded. Since this is so, much more research is needed and many opportunities exist for good, sound investigations.

CHAPTER BIBLIOGRAPHY

1. Silverstein, A. B., "Psychological Testing Practices in State Institutions for the Mentally Retarded," American Journal of Mental Deficiency, LXVIII (1963), 440-445.

BIBLIOGRAPHY

Books

- Cronbach, Lee J., Essentials of Psychological Testing, New York, Harper and Row, 1960.
- Ferguson, George A., Statistical Analysis in Psychology and Education, New York, McGraw-Hill, 1959.
- Goddard, H. H., "A Revision of the Binet Scale," Training School Bulletin, VIII (1911), 56-62.
- Herring, John P., Herring Revision of the Binet-Simon Tests, Yonkers, World Book Co., 1924.
- Hinsie, Leland E. and Robert Jean Campbell, Psychiatric Dictionary, New York, Oxford University Press, 1960.
- Kuhlmann, F., A Handbook of Mental Tests: A Further Revision and Extension of the Binet-Scale, Baltimore, Warwick and York, 1922.
- _____, Tests of Mental Development, Minneapolis, Educational Testing Bureau, 1939.
- Pinneau, Samuel R., Changes in Intelligence Quotient, Boston, Houghton Mifflin Company, 1961.
- Terman, L. M., The Measurement of Intelligence, Boston, Houghton Mifflin Company, 1916.
- _____ and M. A. Merrill, Measuring Intelligence, Boston, Houghton Mifflin Company, 1937.
- _____, Stanford-Binet Intelligence Scale, Boston, Houghton Mifflin Company, 1960.
- Webster's Seventh New Collegiate Dictionary, G. & C. Merriam Company, Publishers, Springfield, Massachusetts, 1963.
- Wechsler, David, The Measurement and Appraisal of Adult Intelligence, Baltimore, The Williams and Wilkins Company, 1958.

_____, Wechsler Intelligence Scale for Children Manual, New York, The Psychological Corporation, 1949.

Yerkes, R. M., J. W. Bridges, and R. S. Hardwick, A Point Scale for Measuring Mental Ability, Baltimore, Warwick and York, 1915.

Articles

Alper, A. E., "An Analysis of the Wechsler Intelligence Scale for Children with Institutionalized Mental Defectives," Dissertation Abstracts, XX (1960), 4711.

Beck, Harry S. and Robert L. Lam, M. D., "The Use of the W.I.S.C. in Predicting Organisity," Journal of Clinical Psychology, XI (1955), 154-158.

Budoff, Milton and Eleanor Mathia Purseglove, "Forms L and L-M of the Stanford-Binet Compared for an Institutionalized Adolescent Mentally Retarded Population," Journal of Clinical Psychology, XIX (1963), 214.

Burnett, A., Comparison of the PPVT, Wechsler-Bellevue, and Stanford-Binet on Educable Retardates," American Journal of Mental Deficiency, LXIX (1965), 712-715.

Carleston, Frederick O. and Chambers L. Stacey, "Evaluation of Selected Short Forms of the WISC," Journal of Clinical Psychology, X (1954), 258-261.

Estes, Betsy Worth, Mary Ellen Curtin, Robert A. DeBurger, and Charlotte Denny, "Relationships Between the 1960 Stanford-Binet, 1937 Stanford-Binet, WISC, Raven, and Draw-A-Man," Journal of Consulting Psychology, XXV (1961), 388-391.

_____, "Relationships Between the Otis, 1960 Stanford-Binet and WISC," Journal of Clinical Psychology, XXI (1965), 296-297.

Finley, Carmen J. and Jack Thompson, "An Abbreviated Wechsler Intelligence Scale for Children for Use with Educable Mentally Retarded," American Journal of Mental Deficiency, LXIII (1958), 473-480.

Guthrie, G. M. and J. J. Pastovic, "Some Evidence of the Validity of the WISC," Journal of Consulting Psychology, XV (1951), 385-386.

- Haynes, Jack R. and S. B. Sells, "Assessment of Organic Brain Damage: A Critical Review," Psychological Bulletin, LX (1963), 316-325.
- Kuhlmann, F., "A Revision of the Binet-Simon System for Measuring the Intelligence of Children," Journal of Psycho-Asthen., Monograph Supplement, I (September, 1912), 41.
- Munpower, Daniel L., "The Fallacy of the Short Form," Journal of Clinical Psychology, XX (1964), 111-113.
- Nale, Stanley, "The Children's Wechsler and the Binet on 104 Mental Defectives at the Polk State School," American Journal of Mental Deficiency, LVI (1951), 419-423.
- Newman, J. R. and F. M. Loos, "Differences Between Verbal and Performance IQs with Mentally Defective Children," Journal of Consulting Psychology, XIX (1955), 16.
- Ogdon, D. P., "WISC IQs for the Mentally Retarded," Journal of Consulting Psychology, XXIV (1960), 187-188.
- Osborne, R. T. and J. Allen, "Validity of Short Forms of the WISC for Mental Retardates," Psychological Reports, XI (1962), 167-170.
- Reger, Roger, "Repeated Measurements with the WISC," Psychological Reports, XI (1962), 418.
- Rohrs, F. W. and M. R. Haworth, "The 1960 Stanford-Binet, WISC and Goodenough Tests with Mentally Retarded Children," American Journal of Mental Deficiency, LXVI (1962), 853-859.
- Sandercock, M. G. and A. J. Butler, "An Analysis of the Performance of Mental Defectives on the Wechsler Intelligence Scale for Children," American Journal of Mental Deficiency, LVII (1952), 100-105.
- Schwartz, L. and E. E. Levitt, "Short-forms of the Wechsler Intelligence Scale for Children in the Educable, Non-institutionalized, Mentally Retarded," Journal of Educational Psychology, LI (1960), 187-190.
- Seashore, H. G., "Differences Between Verbal and Performance IQs on the Wechsler Intelligence Scale for Children," Journal of Consulting Psychology, XV (1951), 62-67.

- Silverstein, A. B., "Effects of Proration on the W.I.S.C. IQs of Mentally Retarded Children," Psychological Reports, XII (1963), 646.
- _____, "Psychological Testing Practices in State Institutions for the Mentally Retarded," American Journal of Mental Deficiency, LXVIII (1963), 440-445.
- _____, "WISC and WAIS IQs for the Mentally Retarded," American Journal of Mental Deficiency, LXVII (1963), 617-618.
- Sloan, W. and B. Schneider, "A Study of the Wechsler Intelligence Scale for Children with Mental Defectives," American Journal of Mental Deficiency, LV (1951), 573-575.
- Stacey, Chalmers L. and Janice Levin, "Correlation Analysis of Scores of Subnormal Subjects on the Stanford-Binet and WISC," American Journal of Mental Deficiency, LV (1951), 590-597.
- Vanderhost, L., W. Sloan, and G. F. Bensberg, "Performance of Mental Defectives on the Wechsler-B llevue and the WISC," American Journal of Mental Deficiency, LVII (1953), 481-483.
- Wright, Clare, "A Modified Procedure for the Abbreviated Revised Stanford-Binet Scale in Determining the Intelligence of Mental Defectives," American Journal of Mental Deficiency, XLVII (1942), 178-184.
- Yalowitz, J. M. and R. G. Armstrong, "Validity of Short Forms of the WISC," Journal of Clinical Psychology, XI (1955), 275-277.

Publications of Learned Organizations

- American Association of Mental Deficiency, A Manual of Terminology and Classification in Mental Retardation, 2nd ed., prepared by R. Heber, Monograph Supplement to the American Journal of Mental Deficiency, 1961.
- Binet, A. and Th. Simon, The Development of Intelligence in Children, Publication #11, Training School, Vineland, New Jersey, May 1916, Williams and Wilkins Co., Baltimore, Maryland, translated by E. S. Kite, from various essays in L'annee psychol.

Unpublished Materials

Poit, Carl H., "Wechsler and Stanford-Binet as Used with the Retardates," a paper presented at the 1965 A. A. M. D. convention in Miami, Florida, 1965.