

A COMPARATIVE EVALUATION OF THE PEABODY  
PICTURE VOCABULARY TEST AS A MEASURE  
OF ABILITY FOR CHILDREN OF  
DIFFERING READING  
PROFICIENCY  
LEVELS

By

JERRY A. LAVITT

Bachelor of Arts  
American International College  
Springfield, Massachusetts  
1953

Master of Education  
Springfield College  
Springfield, Massachusetts  
1959

Specialist in Education  
Springfield College  
1961

Submitted to the Faculty of the Graduate  
School of the Oklahoma State University  
in partial fulfillment of the require-  
ments for the degree of  
DOCTOR OF EDUCATION  
July 28, 1967

JAN 12 1968

A COMPARATIVE EVALUATION OF THE PEABODY  
PICTURE VOCABULARY TEST AS A MEASURE  
OF ABILITY FOR CHILDREN OF  
DIFFERING READING  
PROFICIENCY  
LEVELS

Thesis Approved:

*William B. Ewers*  
\_\_\_\_\_  
Thesis Adviser

*Norman E. Wilson*  
\_\_\_\_\_

*Harry K. Brobst*  
\_\_\_\_\_  
  
\_\_\_\_\_  
  
\_\_\_\_\_

*O. D. Durham*  
\_\_\_\_\_  
Dean of the Graduate School

659323

## ACKNOWLEDGMENTS

Grateful acknowledgment is extended to the writer's committee: Drs. William P. Ewens, Harry K. Brobst, Kenneth A. Browne and Norman E. Wilson.

A special acknowledgment of appreciation is extended to Dr. William P. Ewens, Chairman of the writer's graduate advisory committee for his invaluable encouragement, helpful suggestions and constructive criticisms.

Dr. Edwin Vineyard, the writer's former chairman is acknowledged for his help and assistance.

Special recognition is due Dr. Mary Jane Denton and Mr. Sinai Frenkel for their assistance in administering the tests and their helpful advice.

Sincere gratitude is extended to John, Ina, and Lea Diehm of Stillwater, Oklahoma for their friendship and consideration.

Special recognition is accorded the late Stanley C. Jones, the writer's former principal, for his many past contributions.

## TABLE OF CONTENTS

Chapter	Page
I. INTRODUCTION.....	1
Background of the Problem.....	1
Needs for the Study.....	4
Statement of the Problem.....	5
II. REVIEW OF THE LITERATURE.....	11
Peabody Test Reliability Studies.....	11
Survey of the Peabody Test - Validation Studies.....	15
Comparative Research.....	31
Reading Disability and the Measurement of Intelligence.....	37
III. DESIGN AND METHODOLOGY.....	46
Description of Subjects.....	46
Instruments Used.....	46
Wechsler Intelligence Scale for Children..	47
Tests of Educational Ability.....	48
California Achievement Tests.....	49
Grade Point Average.....	50
Procedure.....	50
IV. RESULTS.....	54
Variability of the I.Q. Test Scores Within Groups.....	54
Correlations Obtained Between the I.Q. Tests.....	56
Predictive Validity of the I.Q. Tests.....	59
Comparisons of I.Q. Intercorrelations.....	60
Comparisons of I.Q. Test Correlations With Grade.....	63
Comparative Analysis of the Differing Levels of Scores Obtained from the I.Q. Tests.....	68
Peabody Test Form A and the Wechsler Test Mean Score Comparisons.....	68
Peabody Test Form A and the TEA Mean Score Comparisons.....	72

Chapter	Page
TEA and Peabody Test Form B Mean Score Comparisons With the Wechsler Test.....	75
V. DISCUSSION.....	80
Discussion of Results.....	80
Limitations.....	84
VI. SUMMARY AND CONCLUSIONS.....	86
Summary.....	86
Conclusions.....	92
BIBLIOGRAPHY.....	95
APPENDICES.....	100

## LIST OF TABLES

Table	Page
I. Correlations Among WISC and PPVT Scores for Each Race - Sex Group.....	27
II. Intragroup Comparisons on I.Q. Scores.....	34
III. Means and Standard Deviations of I.Q.'s.....	55
IV. Intragroup Comparisons of Standard Deviations-Table of F-Levels.....	55
V. Test and Academic Grade Intercorrelations for the Low Reading Ability Group.....	56
VI. Test and Academic Grade Intercorrelations for the Middle Reading Ability Group.....	57
VII. Test and Academic Grade Intercorrelations for the High Reading Ability Group.....	57
VIII. Test and Academic Intercorrelations for All Reading Ability Groups Combined.....	58
IX. PPVT-A and TEA Correlation Comparisons With the WISC Within Each Reading Ability Group....	61
X. PPVT-A and PPVT-B Correlation Comparisons With the WISC Within Each Reading Ability Group....	62
XI. WISC and TEA Correlation Comparisons With Grade Within Each Reading Ability Group.....	63
XII. PPVT-A and TEA Correlation Comparisons With Grade Within Each Reading Ability Group.....	64
XIII. WISC and PPVT-B Correlation Comparisons With Grade Within Each Reading Ability Group.....	65
XIV. PPVT-A and PPVT-B Correlation Comparisons With Grade Within Each Reading Ability Group.....	66
XV. WISC and PPVT-A Correlation Comparisons With Grade Within Each Reading Ability Group.....	67

Table	Page
XVI. Analysis of Variance of Mean Scores Obtained From All Four I.Q. Tests.....	69
XVII. Analysis of Variance of Mean Scores Obtained From Wechsler and Peabody Form A Tests.....	71
XVIII. Analysis of Variance of Mean Scores Obtained From Peabody Form A and TEA Tests.....	73
XIX. Analysis of Variance of Mean Scores Obtained From Peabody Form A and Peabody Form B Tests.....	74
XX. Analysis of Variance of Mean Scores Obtained From Wechsler and TEA Tests.....	76
XXI. Analysis of Variance of Mean Scores Obtained From Wechsler and Peabody Form B Tests.....	78
XXII. Intragroup Comparisons of I.Q. Scores.....	79

## CHAPTER I

### BACKGROUND OF THE PROBLEM

The Peabody Picture Vocabulary Test is an individual intelligence test designed by Lloyd Dunn, a professor at George Peabody College. This test consists of 150 plates, each of which contains four pictures. The pictures on any one plate all represent objects or activities or states of being. The Examiner reads a word and the subject either points at, or orally defines the word by an identifying number. Only ten to fifteen minutes are usually required to give this untimed test and it can be administered only over the critical range of items for a particular subject. The starting point, basal and ceiling vary from testee to testee. The test list includes a book containing the 150 numbered plates and individual test records for the two alternate forms. (The same picture book is used for both forms.) A more complete description of the test can be obtained from the Peabody Picture Vocabulary Test Manual.<sup>1</sup>

In terms of concurrent validity several studies have shown a considerable degree of correlation between the Peabody

---

<sup>1</sup>Peabody Picture Vocabulary Test Manual. Tentative edition. Nashville, Tennessee: George Peabody College, 1958, 31p.



Test and older, widely used, individual measures of intelligence. Varying degrees of correlation between the Peabody Test and group intelligence tests have been reported. In a study of congruent validity between the Peabody Test and Stanford Binet with 315 "educable" children ages 6 to 18 years, the validity coefficient was 0.76.<sup>2</sup>

Since the Peabody is a vocabulary test which does not require reading on the part of the testee, one special utility for it may be with reading handicapped children. One of the problems involved in the assessment of learning ability with reading handicapped children is the limitation of most group tests for this purpose. The score derived from the average group intelligence test is heavily influenced by the reading ability of the testee. The score of any reading handicapped child on a paper and pencil intelligence test must always be highly suspect until further verification is made. Further verification usually consists of the administration of an individual intelligence test. Since the problems in an individual intelligence test are presented orally, the child with poor reading skills is not penalized for this deficiency. The effect of reading disability upon intelligence test performance was shown clearly in a study comparing individual Stanford Binet scores and group-test scores of retarded, normal,

---

<sup>2</sup>Lloyd M. Dunn and Sayde T. Brooks. "Peabody Picture Vocabulary Test Performance and Educable Mentally Retarded Children," Training School Bulletin, 1960, 57. pp. 35-40.

and accelerated readers in the sixth grade.<sup>3</sup> For children whose reading was a year or more accelerated (in relation to Stanford Binet mental age), group-test intelligence scores averaged 15 points higher than than Stanford Binet Intelligence scores. When reading was within plus or minus one year of Stanford Binet mental age, the group test intelligence quotients were two points higher. Where reading was retarded a year or more, group intelligence scores fell 8 points below the Stanford Binet scores. This study points out clearly the caution with which a group intelligence test score must be interpreted for a person who is well below average in reading skills. Thorndike and Hagen state that "a low group test intelligence quotient for a poor reader cannot be taken at face value. It should always be checked with a test that does not involve reading."<sup>4</sup> Non-verbal group tests are sometimes used with reading handicapped children but these usually require some reading and therefore, as Tyler states, "although there are some non-verbal group tests, those most commonly used do penalize non-readers."<sup>5</sup> There is also considerable doubt as to the utility of non-verbal tests in the

---

<sup>3</sup>D. D. Durrell, "The Influence of Reading Ability on Intelligence Measures," Journal of Educational Psychology, 1933, 24. pp. 412-416.

<sup>4</sup>Robert Thorndike and Elizabeth Hagen, Measurement and Evaluation in Psychology and Education. (New York: John Wiley and Sons, Inc., 1961). p. 241.

<sup>5</sup>Leona E. Tyler, Tests and Measurements. (New Jersey: Prentice-Hall, 1963). p. 49.

measurement of academic potential, whereas, vocabulary tests have been found to be highly useful in this area. Terman and Merrill have stated: "We have found the vocabulary test to be the most valuable single test in the Stanford Binet scale."<sup>6</sup>

David Wechsler found the vocabulary sub-test scores to correlate more highly with the Full Scale Intelligence Quotient scores than any other sub-test.<sup>7</sup> Many other studies investigating the measurement of intelligence have shown that vocabulary is the best single item for predicting school success. These findings might appear to support the possibility of the usefulness of the Peabody Test, both with reading handicapped children as well as with those who possess adequate reading skills.

Needs for the Study

Many public school systems are handicapped by a lack of qualified psychometric personnel who can administer individual intelligence tests. In addition to this problem, many guidance counselors find it difficult to find the time needed to administer all of the individual tests which should be given. The fairly common practice of calling in outside psychometric

---

<sup>6</sup>Lewis Terman and Maud Merrill, Measuring Intelligence. (Boston: Houghton Mifflin, 1937). p. 302.

<sup>7</sup>David Wechsler, Wechsler Intelligence Scale for Children Manual. (New York: Psychological Corporation, 1949). p. 100.

personnel to administer individual tests for a school system involves considerable expense and inconvenience. An analysis of the cumulative records of many public school systems often reveals very few, if any, individual tests having been administered. Since group tests of intelligence have many limitations for evaluation of ability with reading handicapped students, and are usually considered to be inferior in validity to individual intelligence tests with normal readers as well, some doubts might be raised as to the adequacy of many school testing programs.

In view of the problems cited above, an individual test which would require no special training and little time to administer would be of great value to public school systems, providing such a test is found to be comparable in validity with the Wechsler or Stanford Binet for use with reading handicapped children. An individual test which can be administered in group form would also be of considerable value for school systems as a general measure of ability for all students if this test were found to be superior to group tests presently being used.

#### Statement of the Problem

The purpose of this study is to compare the individually administered Peabody Picture Vocabulary Test as a measure of ability for students of differing reading proficiency levels with the Tests of Educational Ability, a commonly used group intelligence test requiring reading and a written form of the

Peabody Test which also requires reading on the part of the testee. The Peabody Test will be compared with the other tests in terms of degree of correlation with the Wechsler Intelligence Scale for Children, comparative level of scores obtained, and degree of predictive efficiency of an academic criterion. This study will be expected to contribute information related to the more generally significant question of the value and usefulness of the newer approach to mental assessment through non-reading mediated verbal items, and in particular, the value of this approach with groups of differing reading proficiency. More specifically, this study will attempt to determine if the Peabody Picture Vocabulary Test is superior to a group intelligence test as a measure of ability for students with poorly developed reading skills, and if so, at what level of reading ability will this superiority manifest itself.

The principal hypotheses stated in null form in this study are:

(1) In terms of degree of correlation with the Wechsler Intelligence Scale for children, there will not be a significant difference at the .05 level of confidence between the Peabody Picture Vocabulary Test Scores and the group intelligence test scores within the following groups:

- (1a) the "low" reading ability group
- (1b) the "middle" reading ability group
- (1c) the "high" reading ability group
- (1d) all groups combined

(2) In terms of degree of correlation with the Wechsler Intelligence Scale for Children there will not be a significant difference at the .05 level of confidence between the Peabody Picture Vocabulary Test scores and the written form of the Peabody Test Scores within the following groups:

- (2a) the "low" reading ability group
- (2b) the "middle" reading ability group
- (2c) the "high" reading ability group
- (2d) all groups combined

(3) There will be no significant difference at the .05 level of confidence in the coefficients of correlation obtained between the academic grade point average and the Wechsler Intelligence Scale for Children and group intelligence test within the following groups:

- (3a) the "low" reading ability group
- (3b) the "middle" reading ability group
- (3c) the "high" reading ability group
- (3d) all groups combined

(4) There will be no significant difference at the .05 level of confidence in the coefficients of correlation obtained between the academic grade point average and the Peabody Picture Vocabulary Test and group intelligence test within the following groups:

- (4a) the "low" reading ability group
- (4b) the "middle" reading ability group
- (4c) the "high" reading ability group
- (4d) all groups combined

(5) There will be no significant difference at the .05 level of confidence in the coefficients of correlation obtained between the academic grade point average and the Wechsler Intelligence Scale for Children and the Peabody Test in written form within the following groups:

- (5a) the "low" reading ability group
- (5b) the "middle" reading ability group
- (5c) the "high" reading ability group
- (5d) all groups combined

(6) There will be no significant difference at the .05 level of confidence in the coefficients of correlation obtained between the academic grade point average and the Peabody Picture Vocabulary Test and the Peabody Test in written form within the following groups:

- (6a) the "low" reading ability group
- (6b) the "middle" reading ability group
- (6c) the "high" reading ability group
- (6d) all groups combined

(7) There will be no significant difference at the .05 level of confidence in the coefficients of correlation obtained between the academic grade point average and the Wechsler Intelligence Scale for Children and Peabody Picture Vocabulary Test within the following groups:

- (7a) the "low" reading ability group
- (7b) the "middle" reading ability group
- (7c) the "high" reading ability group
- (7d) all groups combined

(8) There will be no significant difference at the .05 level of confidence between the mean scores obtained from the Peabody Picture Vocabulary Test and the mean of the scores obtained from the Wechsler Intelligence Scale for Children within the following groups:

- (8a) the "low" reading ability group
- (8b) the "middle" reading ability group
- (8c) the "high" reading ability group
- (8d) all groups combined

(9) There will be no significant difference at the .05 level of confidence between the mean scores obtained from the Peabody Picture Vocabulary Test and the mean of the scores obtained from the group intelligence test within the following groups:

- (9a) the "low" reading ability group
- (9b) the "middle" reading ability group
- (9c) the "high" reading ability group
- (9d) all groups combined

(10) There will be no significant difference at the .05 level of confidence between the mean scores obtained from the Peabody Picture Vocabulary Test and the mean of the scores obtained from the Peabody Test in written form within the following groups:

- (10a) the "low" reading ability group
- (10b) the "middle" reading ability group
- (10c) the "high" reading ability group
- (10d) all groups combined



(11) There will be no significant difference at the .05 level of confidence between the mean scores obtained from the Wechsler Intelligence Scale for Children and the mean scores obtained from the group intelligence test within the following groups:

- (11a) the "low" reading ability group
- (11b) the "middle" reading ability group
- (11c) the "high" reading ability group
- (11d) all groups combined

(12) There will be no significant difference at the .05 level of confidence between the mean scores obtained from the Wechsler Intelligence Scale for Children and the mean of the scores obtained from the Peabody Test in written form within the following groups:

- (12a) the "low" reading ability group
- (12b) the "middle" reading ability group
- (12c) the "high" reading ability group
- (12d) all groups combined

## CHAPTER II

### REVIEW OF THE LITERATURE

The review of the literature includes the following areas: (1) a survey of the reliability and validation studies relating to the Peabody Test; (2) comparative research; and (3) reading disability and its relationship to the assessment of intelligence.

#### Peabody Test Reliability Studies

Reliability coefficients for the Peabody Test were obtained by calculating Pearson product moment correlations on the raw scores of the standardization subjects for Forms A and B at each age level. Standard errors of measurement for standard scores (I.Q.'s) were then calculated from the parallel forms reliability coefficient using the formula:  $S.E.M. = \pm \frac{r_a.b.}{\sqrt{2}}$ . Correlations obtained ranged from a low of 0.67 at the six year level to a high of 0.84 at the 17 and 18 year levels with a median of 0.77. The standard errors of measurement of I.Q.'s ranged 6.00 to 8.61, the median being 7.20.<sup>8</sup>

Budoff and Purseglove conducted a reliability study on

---

<sup>8</sup>Peabody Picture Vocabulary Test Manual. Tentative Edition. Nashville, Tennessee: George Peabody College, 1958.

the Peabody Test in which the test (Form A or B) was administered in counterbalanced order to 46 institutionalized 16 to 18 year old retardates. The alternate form was administered one month later. The coefficient of correlation between the two forms was 0.85. Mental age scores were used in the calculations.<sup>9</sup>

Dunn and Brooks administered both forms of the Peabody Test to 371 educable mentally retarded pupils. The students ranged from 6-5 to 18-0 years of age. Forms were given one week apart in counterbalanced order by the same examiner.

The correlation between mental age scores on the two forms was 0.83.<sup>10</sup>

Dunn and Harley administered both forms of the Peabody Test one week apart in counterbalanced order to 20 cerebral palsied children ages 7-1 to 16-2. The coefficient of correlation between the two forms was 0.97.<sup>11</sup> It may be possible that the use of mental age scores tended to inflate the correlation.

Dunn and Hottel conducted a study in which both forms of the Peabody Test were administered one week apart in counter-

---

<sup>9</sup>Milton Budoff and Eleanor Purseglove. "Peabody Picture Vocabulary Test Performance of Institutionalized Mentally Retarded Adolescents," American Journal of Mental Deficiency. 1963. 67, pp. 756-760.

<sup>10</sup>Dunn and Brooks, p. 36.

<sup>11</sup>Randall K. Harley and Lloyd M. Dunn, "Comparability of Peabody, Ammons, VanAlstyne, and Columbia Test Scores with Cerebral Palsied Children," Exceptional Children. 1959, 26. pp. 70-74.

balanced order to 220 trainable retardates. Mental age scores were used in the calculation. The coefficient of correlation obtained between the two forms was 0.84.<sup>12</sup>

Kimball administered both forms of the Peabody Test in counterbalanced order to 62 mentally retarded pupils, ages 10-5 to 15-8. The correlation between the two forms was 0.86.<sup>13</sup>

Moed, Wight, and James conducted a study in which the Peabody Test was re-administered to 29 crippled children after one year in a hospital. The coefficient of correlation between the two administrations was 0.88.<sup>14</sup> This study would appear to indicate considerable temporal stability of Peabody Test scores for children in a restricted environment.

Moss readministered the Peabody Test after a two-year period to 51 educable mentally retarded pupils ranging in age from 6-8 years old. The coefficient of correlation between the two administrations was 0.68.<sup>15</sup> This study also appears

---

<sup>12</sup>Lloyd Dunn and John Hottel, "Peabody Picture Vocabulary Test Performance of Trainable Mentally Retarded Children," American Journal of Mental Deficiency. 1961. 65, pp. 448-452.

<sup>13</sup>Don L. Kimball, "Comparison of Peabody, WISC, and Academic Achievement Scores Among Educable Mental Defectives." Psychological Reports. 1960, 6. p. 502.

<sup>14</sup>George Moed, Byron Wight, and Patricia James, "Inter-test Correlations of the Wechsler Intelligence Scale for Children and Two Picture Vocabulary Tests," Educational and Psychological Measurements. 1963. 23. pp. 359-363.

<sup>15</sup>James Moss, "The Peabody Picture Vocabulary Test with English Children," British Journal of Educational Psychology. 1960. 30. p. 82.

to give some positive evidence regarding the temporal stability of Peabody Test scores.

Shaw obtained a correlation of 0.87 between the two alternate forms of the Peabody Test.<sup>16</sup> Seventy schizophrenics without brain injury were used as subjects in this study. This high correlation may be partially the result of a wide ability range within the group.

Tempero and Ivanoff presented both forms of the Peabody Test to 150 7th grade students utilizing the group administration method. The coefficient of correlation between the I.Q. scores of the two forms was 0.75.<sup>17</sup> This correlation is very comparable to the 0.78 coefficient of correlation reported for this age group on the standardization group.

Hedger conducted a study in which the Peabody Test was administered to 150 orally-trained deaf children ages 6-20 years. Both test forms and mode of presentation were counter-balanced. The modes of presentation were oral and written. In this study the coefficient of correlation obtained between the raw scores of the alternate forms was 0.80.<sup>18</sup> It is possible that the magnitude of the correlation was reduced be-

---

<sup>16</sup>James Shaw, "Comparability of Peabody Test and WAIS Scores with Schizophrenics without Brain Damage," Unpublished Study, Nampa State School, Nampa, Iowa. 1961.

<sup>17</sup>Howard E. Tempero and John M. Ivanoff, "Effectiveness of the Peabody Picture Vocabulary Test with Seventh-Grade Pupils." Unpublished paper. University of Nebraska. 1960.

<sup>18</sup>Mable Hedger, "An Analysis of Three Picture Vocabulary Tests for Use with the Deaf," Research Monogram (In Press). 1964.

cause half of the scores were based on oral, and half on written presentation of the stimulus words.

#### Survey of the Peabody Test Validation Studies

Burnett conducted a study of concurrent validity comparing the Peabody Test with the Wechsler Bellvue and Stanford Binet Tests.<sup>19</sup> The three intelligence tests were administered to 238 residents of a state school for the mentally retarded. Extant scores on the Wechsler Bellvue and Stanford Binet Tests which were obtained over a period of ten years were used. The following results were obtained:

	Correlation
Peabody Test and Wechsler Bellvue Full Score	0.40
Peabody Test and Wechsler Bellvue Verbal	0.47
Peabody Test and Wechsler Bellvue Performance	0.27
Peabody Test and Stanford Binet	0.43

It would seem possible that the use of extant scores obtained over a long period of ten years may have reduced the degree of correlation between the Peabody Test and other measures.

Dunn and Hottel in their study of 220 trainable retardates obtained a coefficient of correlation between the

---

<sup>19</sup>A. Burnett, "Comparison of the Peabody Picture Vocabulary Test with the Wechsler-Bellvue and Stanford Binet on Educable Mentally Retarded Children and Adolescents." Unpublished Study. Owatonna, Minnesota, Owatonna State School, 1964.

Peabody Test and the Stanford Binet of 0.66.<sup>20</sup> In regard to concurrent validity, the Peabody Test was found to correlate 0.39 with reading achievement. Teacher rating scales used to measure the academic accomplishments of the subjects may have reduced the coefficients of correlation.

Himmelstein and Herndon administered Form A of the Peabody Test and the Wechsler Intelligence Scale for Children to 48 children with emotional problems.<sup>21</sup> The children were in the 6-2 to 14-8 year age range. The results were as follows:

	Correlation
Peabody Test I.Q. and Wechsler Full Scale I.Q.	0.63
Peabody Test I.Q. and Wechsler Verbal I.Q.	0.64
Peabody Test I.Q. and Wechsler Performance I.Q.	0.52

Kicklighter administered Form A of the Peabody Test and the Stanford Binet Intelligence Test to 66 Educable mentally retarded children. The testee age range was 6-7 to 16-4. Mental age scores were used in the calculations. The coefficient of correlation between the two tests was 0.87.<sup>22</sup>

Kimbrell conducted a study in which both forms of the Peabody Test, the Wechsler Intelligence Scale for Children,

---

<sup>20</sup>Dunn and Hottel. p. 20.

<sup>21</sup>Philip Himmelstein and James Herndon, "Comparison of PPVT, WISC, and Academic Achievement Scores Among Educable Mental Defectives," Psychological Reports. 1960. 7. p. 502.

<sup>22</sup>R. Kicklighter, "Comparison of PPVT and RSB Test Scores of Educable Mentally Retarded Children," Atlanta, Georgia. State Department of Education, 1964.

and the Gray Votaw-Rogers Achievement Test were administered to 62 retardates in the 10-5 to 15-8 year old age range. I.Q. scores were used in the calculations. The coefficient of correlation between both forms of the Peabody Test and the Wechsler Intelligence Scale for Children (full score) was 0.30. The Peabody Test (Form A) correlated 0.43 with the Wechsler Test Verbal scores. The coefficient of correlation between the Peabody Test and the achievement test scores was only 0.40. The Wechsler Test correlated 0.11 with the achievement scores.<sup>23</sup>

The subjects used in the Kimbrell study were not exposed to a continuous educational program and this may partially explain the very low correlations between the I.Q. tests and achievement scores.

Lindstrom administered Form A of the Peabody Test and the Wechsler Intelligence Scale for Children to 140 children in the kindergarten to 6th grade class range. The coefficient of correlation between the Peabody Test and WISC Full Score I.Q.'s was 0.57. The Peabody Test and WISC Verbal I.Q. correlated 0.67.<sup>24</sup>

A study conducted by Mein using 80 residents in two English institutions for the retarded revealed a coefficient of correlation between the Stanford Binet and Peabody Test

---

<sup>23</sup>Kimbrell, p. 502.

<sup>24</sup>A. Lindstrom, "A Comparison of the Peabody Picture Vocabulary Test and the Wechsler Intelligence Scale for Children," Studies in Minnesota Education. 1961. pp. 131-132.



mental age scores of 0.71.<sup>25</sup>

Another study (unpublished) used a group of students in the 12-14 year old age bracket who ranged in ability from retardates to students who were in advance placement classes.<sup>26</sup>

The validity coefficients were as follows:

	<u>Correlation</u>	<u>Number</u>
Peabody and Wechsler (full scale)	0.82	60
Peabody and Wechsler (verbal)	0.86	60
Peabody and Wechsler (performance)	0.70	60
Peabody and Stanford Binet	0.92	73
Peabody and California Test of Mental Maturity (total score)	0.82	94
Peabody and California Test of Mental Maturity (language score)	0.80	94
Peabody and California Test of Mental Maturity (non-language score)	0.71	94

The author of the Peabody Test has stated that he believes the Peabody and Stanford Binet overlap in the function or functions that they measure.<sup>27</sup> The high degree of correlation shown between the Peabody and Binet in the study above might appear to lend some support to this viewpoint.

Harley and Dunn conducted a study which measured the degree of congruent validity between the Peabody Test and the Revised VanAlstyne Picture Vocabulary Test and the revised Columbia Mental Maturity Scale.<sup>28</sup> Twenty cerebral palsied

---

<sup>25</sup>R. Mein, "Use of the Peabody Picture Vocabulary Test with Severely Subnormal Patients," American Journal of Mental Deficiency. 1962. 67. pp. 296-273.

<sup>26</sup>Dr. Henry Paar and J. Lavitt, "The Peabody Picture Vocabulary Test, A Study of Congruent Validity, Springfield College Psychology Department, Springfield, Massachusetts

<sup>27</sup>Peabody Picture Vocabulary Test Manual. p. 32.

<sup>28</sup>Harley and Dunn. p. 72.

children, ages 7-1 through 16-2 were used as subjects in this study. The Peabody Test correlated .94 with the Revised Van-Alstyne Picture Vocabulary Test, 0.91 with the Ammons Full Range Vocabulary Test, and 0.82 with the Revised Columbia Mental Maturity Scale.

Tempero and Ivanoff correlated Peabody Test Scores for 150 seventh grade children with the Henman-Nelson Tests of Mental Ability and California Tests of Mental Maturity.<sup>29</sup> The validity coefficient was 0.58 with the total California Test scores and 0.61 with the Henmon-Nelson Test.

Several studies are reported of the concurrent validity of the Peabody Test. Tempero and Ivanoff<sup>30</sup> correlated scores for 150 seventh grade children with achievement test scores on the California Achievement Test battery. Correlations ranged from 0.45 to 0.63. Significantly higher relationships were found in the reading areas than in the arithmetic and language mechanics areas. However, all correlations were statistically of high significance.

The study conducted by Harley and Dunn yielded a correlation of 0.90 between Peabody Test scores and teacher ratings of arithmetic achievement, and 0.87 between Peabody scores and teacher ratings of reading achievement.<sup>31</sup>

---

<sup>29</sup>Howard E. Tempero and John M. Ivanoff, "Effectiveness of the Peabody Picture Vocabulary Test with Seventh-Grade Pupils." Unpublished paper. University of Nebraska. 1960.

<sup>30</sup>Tempero and Ivanoff. p. 16.

<sup>31</sup>Harley and Dunn. p. 18.

While the Peabody Test is an individual Test, it can be adapted for administration on a group basis. In a recent study the Peabody Test was administered as a group test by the use of photographic slides of the series of plates. The test was proctored by teachers and one person read off the words. No significant or appreciable differences were found between group and individual administrations.<sup>32</sup>

A study by Richard Weeks examined the use of the Peabody Test in group form with college students.<sup>33</sup> The number of subjects included in this study was 240. Of this number, 50 students were freshmen, 72 students were sophomores, 52 students were juniors, and 66 students were seniors. Both forms A and B of the Peabody Test were administered in group form to the college students. The author reported a reliability coefficient of correlation of 0.87. This compares closely with reliability correlations reported between Peabody forms A and B when the individual method of administration of these tests was used. The author stated that the Peabody Test in group administered form had a high interest value with the tested subjects and was a good rapport establisher. The author of this study also concluded that the Peabody Test was

---

<sup>32</sup>Raymond E. Norris, John V. Hottel, and Sayde Brooks, "Comparability of Peabody Picture Vocabulary Test Scores Under Group and Individual Administration," Journal of Educational Psychology. 1960, 51. pp. 87-91.

<sup>33</sup>Richard Weeks, "Effectiveness of the Peabody Picture Vocabulary Test with College Students," Journal of Educational Research. 1963, 57. p. 131.

effective for obtaining a verbal intelligence screening for college students. He based this conclusion on the fact that the subjects scored much above the established norms for 18 year olds.

Moss conducted a study in which the Peabody Test, Stanford Binet, Primary Mental Abilities Test, and selected achievement tests were administered to 51 seven year old educable mentally retarded children. Mental age scores were used in the calculations. Results regarding concurrent validity were as follows:<sup>34</sup>

	Correlation
Peabody Test and Stanford Binet	0.60
Peabody Test and Primary Mental Abilities Test	0.82
Stanford Binet and Primary Mental Abilities Test	0.56

Coefficients of correlation regarding concurrent validity of the intelligence tests were as follows:

	Correlation
Peabody Test and Metropolitan Reading Test	0.32
Stanford Binet and Metropolitan Reading Test	0.51
Peabody Test and Metropolitan Arithmetic Test	0.51
Stanford Binet and Metropolitan Arithmetic Test	0.70
Peabody Test and General Information Test	0.68
Stanford Binet and General Information Test	0.72

---

<sup>34</sup>James W. Moss, "An Evaluation of the Peabody Picture Vocabulary Test with the PMA and 1937 Stanford Binet on Trainable Children." Unpublished paper. Urbana, Illinois. University of Illinois. Institute for Research on Exceptional Children. 1962.

Predictive correlations were also reported in this study which indicated the Primary Mental Abilities Test excelled the Peabody Test in predicting school achievement:

<u>1960 Intelligence Tests</u>	<u>1961 Achievement Tests</u>	
	<u>Reading</u>	<u>Arithmetic</u>
Peabody Test	.22	.43
Stanford Binet	.44	.68
Primary Mental Abilities Test	.49	.78

Moss and Edmonds administered the Peabody Test and Otis Intelligence Test to 101 English children in the 6-3 to 8-5 range. Mental age scores were used in the calculations. The coefficient of correlation between the two tests was 0.68.<sup>35</sup>

Saslow conducted a study in which the Peabody Test and Stanford Binet were administered to 37 children, ages 2-8 to 13-8, who were known to have, or suspected of having cerebral palsy. The coefficient of correlation obtained between the two intelligence tests was 0.82.<sup>36</sup>

Saslow and Larsen administered the Peabody Test, Children's Picture Test, Stanford Binet, and Vineland Test to 31 children, ages 5-7 years, who were attending an out-patient clinic for crippled children. The following results were

---

<sup>35</sup>Moss and Edmonds. p. 64.

<sup>36</sup>H. Saslow, "The Comparability of the Peabody Picture Vocabulary Test and Revised Stanford Binet, Form L-M, with Cerebral Palsied Children," Paper read at the American Psychological Association Meeting, New York City, New York. August, 1961.

obtained.<sup>37</sup>

	Correlation
Peabody Test and Stanford Binet	0.71
Peabody Test and Vineland Test	0.43
Peabody Test and Children's Picture Test	0.63

A study conducted by Shaw compared the Peabody Test and Wechsler Adult Intelligence Scale with 70 schizophrenics without brain damage. The following coefficients of correlation were obtained:<sup>38</sup>

	Correlation
Peabody Test Form A and WAIS Full Score	0.79
Peabody Test Form B and WAIS Full Score	0.83
Peabody Test Form A and WAIS Verbal	0.86
Peabody Test Form B and WAIS Verbal	0.82
Peabody Test Form A and WAIS Performance	0.62
Peabody Test Form B and WAIS Performance	0.63
Peabody Test Form A and WAIS Vocabulary	0.81
Peabody Test Form B and WAIS Vocabulary	0.85

Tobias and Gorelick administered the Peabody Test and Wechsler Adult Intelligence Scale to 107 retarded adults, ages 17 into the 30's. The coefficient of correlation between the I.Q. scores of the two tests was 0.61. Extant

---

<sup>37</sup>H. Saslow and E. Larsen, "The Comparability of the Peabody and Children's Picture Tests, Stanford Binet, and Vineland Scales with Cerebral Palsied Children." Paper read at Rocky Mountain Psychological Association Meeting, 1963.

<sup>38</sup>Shaw. p. 27.

Stanford Binet scores were available for 69 of the subjects. The Binet correlation with the Peabody Test was 0.69. The coefficient of correlation between the Peabody Test and Wide Range Reading Achievement Test scores was 0.52.<sup>39</sup>

Wolfenberger conducted a study in which the Peabody Test and Wide Range Achievement Test were administered to 61 institutionalized retardates, ages 11-43. Peabody Test mental age scores correlated 0.52 with reading, 0.33 with spelling, and 0.35 with arithmetic.<sup>40</sup> These values are appreciably above those found by Kimbrell and are comparable with those found by Tobias and Gorelick. It is possible that the type of schooling received by institutionalized retardates may reduce the degree of correlation with the Peabody Test.

Budoff and Purselove in their study of 238 residents of a state school for the mentally retarded obtained the following correlations between the Peabody Test and the Stanford Binet Intelligence Test:<sup>41</sup>

---

<sup>39</sup>Jack Tobias and Jack Gorelick, "The Validity of the Peabody Test as a Measure of Intelligence of Retarded Adults." Training School Bulletin. 1961. 50. pp. 92-98.

<sup>40</sup>Wolf Wolfenberger, "The Correlation Between PPVT and Achievement Scores among Retardates: A Further Study." American Journal of Mental Deficiency. 1962. 67. pp. 450-454.

<sup>41</sup>Budoff and Purselove. p. 759.

	Correlation
Peabody Test Form A and Stanford Binet Form L	0.85
Peabody Test Form A and Stanford Binet Form LM	0.88
Peabody Test Form B and Stanford Binet Form L	0.86
Peabody Test Form B and Stanford Binet Form LM	0.83

The correlations obtained in this research appear to concur closely with the findings of Lavitt, Kicklighter, Mein, and Moss and Edmunds studies cited earlier in this report.

Corwin conducted a study using two groups of grade 4, 5, and 6 children -- one of Mexican descent (Spanish-speaking background) and the other of Anglo-Saxon descent (English-speaking background). The Peabody Test and Wechsler Intelligence Scale for Children were administered among other tests.<sup>42</sup> The coefficients of correlation reported between the two tests were 0.52 for the Mexican group and 0.61 for the Anglo-Saxon group.

Moed, Wright and James obtained a coefficient of correlation of 0.84 between the Peabody Test and the Wechsler Intelligence Scale for Children.<sup>43</sup> Eighty-three crippled children were the subject of this research.

---

<sup>42</sup>Betty J. Corwin, "The Influence of Culture and Language on Performance on Individual Ability Tests," Unpublished Study. Northridge, California. San Fernando Valley State College. 1962.

<sup>43</sup>Moed, Wright and James. p. 351.



Garret conducted a study with 50 sixth-grade students who were labeled "brighter-than-average." The coefficient of correlation obtained between the Peabody Test and the Co-operative School and College Ability Test verbal score was 0.75. The degree of correlation reported between the Peabody Test and Sequential Test of Educational Progress Reading Test was 0.50.<sup>44</sup>

A study by Robert Hughes and Kenneth Tessler was conducted to determine if the Peabody Test could appropriately be substituted for the WISC as an individual test of intelligence for Negro and White culturally deprived children. All of the subjects utilized in this study were suspected of being mentally retarded. The subjects were examined individually by white examiners. The WISC and the PPVT (Form A) were alternated to preclude any possible order effects.

The data for this study consisted of the verbal, performance and full-scale I.Q. scores from the WISC and the PPVT I.Q. score. Correlations among these scores were computed. In order to test for the linear effects of age as well as race, sex and race-sex interactions an analysis of variance was used in the analysis of data.

The correlations among the test scores are shown in Table I. The correlations between all WISC verbal and full-

---

<sup>44</sup>Jane Garrett, "Comparisons of the Peabody Picture Vocabulary Test and Wechsler Intelligence Scale for Children." Unpublished Study. Nashville, Tennessee. George Peabody College for Teachers. 1959.

(From "Comparison of WISC and Peabody Scores of Negro and White Rural Children," by Robert Hughes and Kenneth Tessler.)

Table I

Correlations Among WISC and PPVT Scores  
For Each Race-Sex Group

<u>Negro Male (n=49)</u>				<u>Negro Female (n=34)</u>			
WISC Verb.	WISC Perf.	WISC Full	PPVT	WISC Verb.	WISC Perf.	WISC Full	PPVT
1.00	0.58**	0.85**	0.61**	1.00	0.41*	0.74**	0.56**
..	1.00	0.92**	0.58**	..	1.00	0.89**	0.41*
..	..	1.00	0.66**	..	..	1.00	0.56**
..	..	..	1.00	..	..	..	1.00
<u>White Male (n=32)</u>				<u>White Female (n=22)</u>			
WISC Verb.	WISC Perf.	WISC Full	PPVT	WISC Verb.	WISC Perf.	WISC Full	PPVT
1.00	0.35*	0.79**	0.43*	1.00	0.50*	0.86**	0.42*
..	1.00	0.84**	0.21	..	1.00	0.87**	0.56**
..	..	1.00	0.38*	..	..	1.00	0.55**
..	..	..	1.00	..	..	..	1.00
*p	.05	**	.01				

scale I.Q.'s and the PPVT scores were statistically different from zero in the positive direction.

Despite relatively large standard errors of estimate, the authors conclude that the PPVT has a distinct advantage over group tests of intelligence for the culturally deprived.<sup>45</sup>

Klaus and Starke conducted a study involving the predictive validity of the Peabody Test. The Peabody Test was administered to 270 beginning Grade 1 children. The following spring the children were given the Metropolitan Achievement Test, Primary 1 Battery. Coefficients of correlation reported were as follows:<sup>46</sup>

	Correlation
Peabody Test and Word Knowledge	0.39
Peabody Test and Word Discrimination	0.35
Peabody Test and Reading	0.39

While statistically significant, the predictive validity coefficients reported in the Klaus Starke study fall in the low range.

Allan, Haupt and Jones compared the Peabody Test and Wechsler Intelligence Scale for Children using two groups

---

<sup>45</sup>R. Hughes and K. Tessler, "Comparison of WISC and Peabody Scores of Negro and White Rural Children." American Journal of Mental Deficiency. 1965. 69. pp. 877-880.

<sup>46</sup>R. Klaus and Christiana Starke, "Experimental Revision of the Peabody Picture Vocabulary Test as a Predictor of First Grade Reading Ability." Unpublished Study. Nashville, Tennessee. Peabody College, Psychology Department. 1964.

that differed in perceptual ability.<sup>47</sup> One group was composed of 20 educable mentally retarded children who scored well on the Frostig Developmental Test of Visual Perception. The other group was composed of 20 low scoring subjects of the same classification and age. The mean WISC I.Q. was 75.4 for the high perceivers and 54.8 for the low perceivers. The mean Peabody Test I.Q. was 81.7 for the high and 70.1 for the low perceivers. The authors suggest that the Vocabulary Test overestimates the intellectual efficiency of the low perceivers. A probable reason that the Peabody Test I.Q. differential was less than for the WISC is that the former is measuring hearing vocabulary with a minimum of visual skills needed.

Mueller presented alternate forms of the Peabody Test with plates reproduced in regular and enlarged sizes to 39 visually limited children. Children with visual acuity 10/200 to 20/200 performed better on the large plates, with no difference for pupils in the 20/20 to 20/200 visual acuity range.<sup>48</sup> Apparently children classified as partially seeing may use the regular Peabody test plates effectively.

Shipe, Cranwell and Dunn conducted a study in which groups of withdrawn, acting out, and non-disturbed retardates

---

<sup>47</sup>R. M. Allen and R. W. Jones, "A Suggested Use and Non-Use for the Peabody Picture Vocabulary Test with the Retarded Child," Psychological Reports. 1964. 15. pp. 421-422.

<sup>48</sup>Max Mueller, "Effects of Illustration Size on PPVT Test Performance of Visually Limited Children," Exceptional Child. 1962. 29. pp. 124-128.

in residential settings of 20 each were given the Peabody Test for children. Coefficients of correlation were reported as follows:<sup>49</sup>

	<u>Peabody Test</u>	<u>WISC</u>
Withdrawn	0.60	0.48
Acting Out	0.64	0.52
Non-disturbed	0.67	0.52

A review of the studies regarding the concurrent validity of the Peabody Test reveals a considerable degree of correlation between this test and older, well-established, individual measures of ability.

One limitation regarding the research findings is that only one validation study concerned itself with the utility of the Peabody Picture Vocabulary Test for subjects with reading difficulties. This study, conducted by Donald Neville, is described in Chapter III, Review of Literature, in the Comparative Research section of this work.

Comparatively few of the validation studies attempted to procure data dealing with the predictive validity of the Peabody Test. Further, few of the studies utilized academic achievement as measured by school grades in addition to achievement test scores as the criterion for concurrent validity.

---

<sup>49</sup>Dorothy Shipe, R. Cromwell, and L. Dunn, "Responses of Emotionally Disturbed and Non-Disturbed Retardates to PPVT Items of Human vs. Non-Human Content," Journal of Consulting Psychology. (Submitted), 1964.

Many of the Peabody Test validation studies utilized atypical subjects. Many of these subjects were either mental retardates or persons attending schools for exceptional children. This could conceivably decrease the value of the research findings as to use with more normal groups.

#### Comparative Research

Donald Neville conducted a research study in 1964 dealing with the utility of the Peabody Test as a valid measure of ability for poor readers.<sup>50</sup>

Neville's study sets out to answer the following three questions:

- (1) Does the lack of reading ability negatively influence scores of verbally oriented group I.Q. tests for pupils in the fifth grade:
- (2) At what level does lack of reading ability influence the intelligence test results to a degree which would invalidate them?
- (3) Could a short, easily administered test of intelligence, the Peabody Test, neutralize the influence of low reading ability to the same degree as a longer, more difficult to administer individual test like the Wechsler Intelligence Scale for Children?

---

<sup>50</sup> Donald Neville, "The Relationship Between Reading Skills and I.Q. Test Scores." Unpublished Study. Nashville, Tennessee. George Peabody College for Teachers. 1964.

Neville selected 148 fifth graders in two urban schools located in upper, lower and middle class neighborhoods. The pool of subjects were divided into three groups according to reading achievement as measured by the Metropolitan Achievement Test. The Metropolitan Test was administered during the last two months of the previous school year. Those whose scores were below 4.00 grade level were assigned a group labeled "poor readers." Those whose scores were 4.00 through 4.99 grade levels were designated as "average readers," and those scoring above 4.99 grade level were labeled as "good readers." Since the poor reading group contained 20 subjects, 14 males and six females were randomly selected for the other two groups. Neville states that the groups were equated as to sex because of the influence of this variable on Wechsler Intelligence Scale for Children scores. The Wechsler Test and the Peabody Picture Vocabulary Test were then administered to each subject. After the individual tests were administered, data was available on only 18 of the selected 20 poor readers. The two subjects on whom data was unavailable were both boys. In view of this fact, two males were randomly omitted from the other groups. Thus, Neville's results were based on three groups of 18 subjects each.

The statistical plan in Neville's study for handling the data included analysis of variance and correlational techniques. The primary use of the analysis of variance approach was to answer the question of whether or not lack of reading ability significantly influences I.Q. scores on a verbal-type

group test and the degree of reading deficiency necessary to significantly lower the group scores. To accomplish this, the three reading ability groups were compared as to I.Q. scores achieved on five tests: (1) Lorge Thorndike; (2) Wechsler Intelligence Scale for Children, Verbal, Performance, and Full Scale I.Q.'s; (3) The Peabody Picture Vocabulary Test. The comparison was accomplished by the use of a mixed analysis of variance design. The correlation technique was employed to examine the relationship between the Wechsler and Peabody Test scores, in order to discover whether or not the Peabody Test could be used as a substitute for the Wechsler Test.

In order to answer the first question relating to the influence of I.Q. scores on a verbally oriented group test several analyses of variance were completed. The three reading groups were treated as levels while the various I.Q. scores were conceived of as treatments.

The first analysis compared the three reading levels and all five I.Q. scores and resulted in F ratios significant beyond the .01 level on the between levels (reading groups) and interaction comparisons. The difference between tests were not significant. A visual inspection of a graphic representation of the group means on the various tests led Neville to believe that the differences between the Lorge Thorndike and the other tests were the primary factor which contributed to the significant F ratios.

Neville's next step was to complete four more Type I



analyses comparing the three reading levels Large Thorndike I.Q. scores with each of the other four I.Q. scores. When significant interaction and between F ratios resulted, tests were applied to determine specifically where the significant differences occurred.

TABLE II

INTRAGROUP COMPARISONS ON I.Q. SCORES<sup>1</sup>

Comparison	Differences Between Means		
	PR	AR	GR
WISC-V -LT	7.56**	.50 (NS)	-4.06
WISC-P -LT	11.73**	1.17 (NS)	-7.48**
WISC-PS -LT	8.00**	.39 (NS)	-5.06*
PPVT - LT	11.28**	1.42 (NS)	3.72 (NS)

(From the "Relationship Between Reading Skills and Group Test Scores" by Donald Neville.)

<sup>1</sup>To interpret Table II: In the first row differences were computed by subtracting mean LT I.Q.'s from mean WISC-V I.Q.'s. For the PR group the mean WISC-V I.Q. was 7.46 points higher, for the GR group the mean WISC-V I.Q. was 4.06 points lower.

\*significant at .05 level

\*\*significant at .01 level

Neville's summary of the tests reveals that the "good readers" scored significantly higher than the "poor readers" or "average readers" on all measures. The "average readers" had a mean score significantly higher than that of the "poor

readers" on the Lorge Thorndike, Wechsler Verbal Score and Wechsler Full Scale Score. However, these two groups were not significantly different on the Wechsler Test Performance and Peabody Test measures.

An intragroup comparison of the performance on the Lorge Thorndike and each of the other measures was conducted. It was discovered that the "poor reading" group's performance on the Lorge Thorndike was always significantly lower than its performance on the individual tests. The "average reading" group made scores on the individual tests all of which were lower than those made on the Lorge Thorndike.

Neville lists six conclusions in regard to the above findings. They are as follows:

(1) "Poor readers," in the middle elementary grades tended to make scores on group I.Q. tests requiring reading which were significantly lower than those scores made on individual tests requiring little or no reading.

(2) "Good readers," in the middle elementary grades, tended to make scores on the group I.Q. tests which were as high as or higher than their scores on the individual tests.

(3) "Average readers," in the middle elementary grades, tended to make scores on most I.Q. measures which were not different from their scores on individual tests.

(4) "Poor readers" tended to make scores on most I.Q. measures which were significantly lower than those made by "good readers."

(5) "Poor readers" were found to make I.Q. scores sig-

nificantly lower than "average readers" on the Lorge Thorndike, Wechsler Verbal, and Wechsler Full Score, but not significantly different on other tests (Wechsler Performance and the Peabody Test).

(6) "Good readers" made I.Q. scores on all measures which were significantly superior to "poor" or "average" readers.

Neville concludes that reading ability does tend to negatively influence scores on verbally oriented group I.Q. tests for pupils in grade five. He further concludes that it appears that a 4.0 achievement level in reading is a minimum for obtaining reasonably valid I.Q. scores for children in intermediate grades. This conclusion is based on two sets of data. First, the "average" and "poor" reading groups were different on the Lorge Thorndike Test but not different on most of the individual measures. Further, the "average" group had Lorge Thorndike scores not different from their individual I.Q. scores while the "poor" group had Lorge Thorndike scores significantly lower than individual test scores. Since the "average" reading group had reading levels between 4.00 and 4.99 grade level, Neville concludes that 4.00 grade level is minimal if one is to put reliance in a verbal type I.Q. test in the middle grades.

Neville's results show that 66.6 percent of the "poor reading" group had Wechsler scores six or more points higher than their scores on the Lorge Thorndike. Only 27.7 percent of the "average reading" group and 16.6 percent of the "good

readers" fell into this category. The differences between the Lorge Thorndike and Peabody Test followed the same pattern.

Neville used two major analyses to explore the relationship between the Peabody Test and Wechsler Intelligence Scale for Children. First, a t-test was applied to ascertain whether or not the mean Peabody Test I.Q. of the total group (54 subjects) was significantly different from the mean Wechsler Test I.Q. The scores were found not to differ significantly.

Secondly, product moment correlations between the I.Q.'s of the Peabody and Wechsler Test were computed for each of the three groups. The resulting coefficients of correlation for each group were found to be as follows:

	<u>Correlation</u>
Peabody Test and Wechsler (good readers)	0.42
Peabody Test and Wechsler (average readers)	0.65
Peabody Test and Wechsler (poor readers)	0.66

Neville states that the results of this study would indicate that the Peabody Test can serve as an adequate substitute for the Wechsler Intelligence Scale for Children. Apparently this conclusion is limited in scope to include intermediate grade children.

Reading Disability and the Measurement  
of Intelligence

Many reading researchers appear to concur in the belief that lack of reading ability creates serious problems in regard to obtaining a valid intellectual assessment with group intelligence tests. Harris states that many commonly used group intelligence tests are unsuitable for use with poor readers because their questions are presented in printed form. A child with average, or even superior intelligence, has difficulty in reading the questions. This author considers such widely used group tests as the Otis Self-Administering I.Q. Test, Henmon-Nelson I.Q. Test, and the Army Alpha Test as relatively useless in the study of poor readers.<sup>51</sup>

Strang expresses the belief that most group intelligence tests are not very useful in the analysis of poor readers because of their primarily verbal composition. Most of the group intelligence tests require reading skills, especially reading comprehension.<sup>52</sup>

Wheeler cautions that any individual who fails to develop a reading proficiency level commensurate with his mental ability will be handicapped when given an intelligence test re-

---

<sup>51</sup>Albert J. Harris, How To Increase Reading Ability (New York: Longmans, Green and Company, 1961). p. 224.

<sup>52</sup>Ruth Strang, Problems in the Improvement of Reading in High School and College (New York: McGraw Hill Book Company, 1946). p. 210.

quiring reading beyond his reading achievement level.<sup>53</sup>

Cronbach expresses the idea that intelligence tests which are strictly verbal in content will yield scores that are greatly influenced by the testees' reading ability and familiarity with the language. This writer points out that it is dangerous to interpret a poor reader's low score on these tests as a sign of deficient mental ability.<sup>54</sup>

Durrell suggests that intelligence tests with a large number of reading items should be labeled reading tests because they measure this achievement factor more accurately than anything else.<sup>55</sup>

While many of the reading and testing writers express the apparent logically sound view that poor readers are penalized on group intelligence tests requiring reading, there is a considerable amount of conflicting evidence. Blair and Kaman conducted a study to determine whether a typical group intelligence test requiring reading ability (Otis Self-Administering Test of Mental Ability) gives disproportionately low scores to college freshmen who are poor readers. A test not requiring reading (Revised Beta Examination) was administered to the students for comparative purpose. Two

---

<sup>53</sup>Lester R. Wheeler, "The Relation of Reading to Intelligence," School and Society. 1949. 70. pp. 225-227.

<sup>54</sup>Lee J. Cronbach, Essentials of Psychological Testing. 2d. ed. (New York: Harper Brothers, 1960). p. 220.

<sup>55</sup>Durrell. p. 221.

groups of college freshmen were selected. One group labeled "poor readers" had an average percentile rank in the Iowa Silent Reading Test of 13. The second group labeled "good readers," had an average percentile score of 91 on the same reading test. The results of this study show that both the "good readers" and "poor readers" made relatively higher scores on the Otis Test which requires reading, than they did on the Revised Beta Examination where reading skill is not required. The authors conclude that the Otis Test does not underestimate the intelligence of poor readers at the college freshmen level. It was further concluded that there is a need for more research of this type to determine how far down the educational ladder one must go before reading involved on a test of the Otis type begins to discriminate against poor readers.<sup>56</sup>

Another study at the college level which was conducted with 1681 freshmen, found a high degree of relationship existing between the linguistic scores on the ACE Psychological Examination and reading ability. Conversely, a low degree of relationship was found to exist between quantitative scores on the ACE and reading. The authors concluded that the ACE Psychological examination is materially influenced by reading efficiency, and special consideration should be given this

---

<sup>56</sup>Glenn Blair and James Kaman, "Do Intelligence Tests Requiring Reading Ability Give Spuriously Low Scores to Poor Readers at the College Freshmen Level?", Journal of Educational Research. 1942. 36. pp. 280-283.

factor in interpretation and use of the AGE.<sup>57</sup>

Stake and Mehrens attempted to determine if "poor readers" at the sixth grade level were handicapped in taking a group intelligence test requiring reading. In this study the investigators wanted to learn if "retarded readers" would score higher on an individual intelligence test than would be predicted from the results of a group test requiring reading. Reading ability was measured with the California Aptitude Test. The California Test of Mental Maturity (CTMM) served as the group test requiring reading. The Wechsler Intelligence Scale for Children was the individual test utilized. Thirty-one "retarded readers," 31 "accelerated readers," and 11 "normal readers" were chosen as subjects. For each individual a Wechsler Intelligence Scale for Children mental age was predicted from his California Test (CTMM) mental age. From this predicted Wechsler Test mental age the actually measured Wechsler Test mental age was subtracted. The mean differences for the "retarded readers" were compared to those of the "accelerated" and "normal readers." The observed mental ages for the "retarded readers" were essentially the same as those predicted by the group test. The authors conclude that the "retarded readers" in this study were not handicapped in taking a group intelligence test re-

---

<sup>57</sup>Lester Wheeler and Viola Wheeler, "The Relationship Between Reading Ability and Intelligence Among University Freshmen," Journal of Educational Psychology. 1960. 51. pp. 230-237.



quiring reading.<sup>58</sup> It would seem that one possible weakness associated with this study is the small size of the "normal reading" group used.

Bliesmer conducted a study with the purpose of finding out to what extent results obtained from various group intelligence tests would approximate results obtained from an individual test for reading handicapped children. The Kuhlmann-Anderson Test and California Short Form Test of Mental Maturity were compared with the Stanford-Binet as to results with 80 poor readers. The grade placement of these children ranged from grade 4 to grade 7. It was discovered that neither of the group tests yielded estimates which might be considered adequate approximations of Stanford Binet estimates.<sup>59</sup> If this study had included data for average or above average readers the significance of the findings might possibly have been more clearly defined.

One study conducted with 309 7th grade students compared the I.Q.'s obtained on verbal and non-verbal group tests of poor readers and normal readers. The authors examined "discrepancy scores" (excess of non-verbal I.Q. over verbal I.Q.) and hypothesized that if reading ability is related to the verbal score but not to the non-verbal score, there should be

---

<sup>58</sup>Robert Stake and William Merens, "Reading Retardation and Group Intelligence Test Performance," Exceptional Children. May, 1960. 26. pp. 497-501.

<sup>59</sup>Emery Bliesmer, "A Comparison of Results of Various Capacity Tests Used With Retarded Readers," Elementary School Journal. 1956. 56. pp. 400-402.

a significant difference between the mean discrepancy score of "retarded readers" and the mean discrepancy score of "non-retarded readers." Two hundred and sixty-six students labeled "retarded readers" and 43 students labeled "non-retarded readers" were used in this study. The Pintner Verbal and Pintner Non-Language Tests were administered to all of the students in this study. In general, the "non-retarded" reading group received lower scores on the Pintner Non-Language Test while the "retarded readers" had lower scores on the verbal section. Significant differences were found to exist between the discrepancy scores of the two groups. The authors concluded that low intelligence quotients obtained by poor readers may reflect their reading retardation rather than a basic inability to learn. The authors express the belief that the use of group tests such as the Pintner Verbal Test gives an erroneous picture of the learning capacity of poor readers.<sup>60</sup>

Barbe and Griek found a high degree of correlation existing between the total reading score on the Iowa Silent Reading Test and I.Q. as measured by the Henman-Nelson Intelligence Test. The correlation existing between reading rate and the Henman-Nelson I.Q. was not statistically significant. The authors cite the need for more research in this

---

<sup>60</sup> Emma Platter, Stanton Platter, Clarence Sherwood, and Sylvia Sherwood, "Relationship Between Reading Retardation and the Measurement of Intelligence," Personnel and Guidance Journal. September, 1959. 38. pp. 49-51.

61  
area.

Manolakes and Sheldon conducted a study to determine what effect the reading ability of a pupil would have upon his intelligence quotient as measured by the California Test of Mental Maturity (language-factors section). The language-factors intelligence quotient was selected because it contributed to the total intelligence quotient of the test and is highly verbal in content. The authors attempted to discover at what grade level reading ability would be most related. Although the results of this study were inconclusive, some evidence was revealed that indicated high correlations exist between reading ability and I.Q. (as measured by this test) above the grade 5 level. However, this was not consistent for every school. The authors speculate that a "fourth-grade hump" in reading may actually exist as postulated by White.<sup>62</sup> During the period of transition from the primary to the intermediate grade level, many changes took place in the reading skills required. These changes may cause a leveling-off in the reading powers of individuals who have gained the primary skills but lack sufficient reading ability to cope with the demands of a new situation. Therefore, if this hypothesis is correct, group intelligence tests

---

<sup>61</sup>Walter Barbe and Werner Griek, "Correlations Between Reading Factors and I.Q.," School and Society, March, 1962. 72. pp. 134-136.

<sup>62</sup>Margaret L. White, "Eliminating the Fourth Grade Hump in Reading," Monograph on Language Arts. 1948. 59. p. 112.

may penalize poor readers in the primary grades to a lesser degree than for the high grades. This assumption is in agreement with the findings of Neville who states that a 4.0 grade level is necessary for a valid assessment of intelligence with the Lorge Thorndike Test.<sup>63</sup> Manolakes and Sheldon conclude that more research is necessary in this area.<sup>64</sup>

In summation, while there is evidence to support the conclusion that reading ability affects performance on group intelligence tests requiring reading, a number of factors appear to affect this relationship. Among these factors are: (1) reading level of the student, (2) type of group test used, and (3) grade level. Present research evidence does not give conclusive answers, especially as to how much reading deficiency is necessary to lower significantly the I.Q.'s obtained from the different group tests.<sup>65</sup>

---

<sup>63</sup>Neville. p. 7.

<sup>64</sup>George Manolakes and William Sheldon, "The Relation Between Reading Test Scores and Language Factors Intelligence Quotients," Elementary School Journal. February, 1955. 55. pp. 347-350.

<sup>65</sup>James B. Stroud, "A Note on the Relationship Between Reading and Intelligence Scores," Research and Evaluation in College Reading. Ninth Yearbook National Reading Conference For College and Adults. 1960.

## CHAPTER III

### DESIGN AND METHODOLOGY

The design and methodology of the study is presented in this chapter according to the following topics:

(1) Description of the Subjects, (2) Instruments Used, and (3) Procedure.

#### Description of Subjects

One hundred and twenty-six seventh grade students of Stillwater Junior High School, Stillwater, Oklahoma, were selected as subjects. The mean chronological age of these students was 13-0 when they were selected for this study in December, 1963. There were 62 girls and 64 boys in the group. From the original group of 126 students, only 113 (57 boys and 56 girls) remained as subjects after random withdrawals were made to obtain homogeneity of variance of I.Q. scores within the three reading ability groups.

#### Instruments Used

In this study four intelligence tests were used. These were: (1) Peabody Picture Vocabulary Test (Form A), (2) Peabody Picture Vocabulary Test (Form B), (3) Wechsler Intelligence Scale for Children, and (4) Tests of Educational Ability.

The California Achievement Tests (Total Reading Section) and a grade point academic average were also utilized.

#### Wechsler Intelligence Scale for Children

The Wechsler Intelligence Scale for Children was used in this study as a standard against which the other intelligence tests were to be compared as to results within the three different reading ability groups.

Correlation coefficients between the Stanford-Binet and WISC Full Scale I.Q.'s have been found to vary from 0.75 to 0.90; for the Verbal-Scale, between 0.65 and 0.90; and the Performance Scale has been found to correlate closely with the Arthur Performance Scale.<sup>66</sup>

Freeman concludes that on the basis of research thus far reported, the WISC Full Scale and Verbal Scale intelligence quotients share considerable communality of psychological functions being measured with the Stanford Binet I.Q.<sup>67</sup>

Anastasi states that the establishment of norms and the determination of reliability for the WISC are models of good test construction.<sup>68</sup>

---

<sup>66</sup>J. J. Pastovic and G. M. Guthrie, "Some Evidence on the Validity of the WISC," Journal of Consulting Psychology. 1951. 15. pp. 385-386.

<sup>67</sup>Frank S. Freeman, Theory and Practice of Psychological Testing (New York: Holt Rinehart and Winston, 1962). p. 274.

<sup>68</sup>Anne Anastasi, Psychological Testing (New York: The Macmillan Company, 1957). p. 324.

A more complete description of the WISC can be obtained in Freeman's text<sup>69</sup> or in the Wechsler Test Manual.<sup>70</sup>

#### Test of Educational Ability

The SRA Tests of Educational Ability (TEA) was used in this study to represent a typical group intelligence test requiring reading which would be compared as to scores obtained within the three reading ability groups with the WISC and Peabody Test (Form A) scores. The TEA are designed to provide three aptitude measures for judging a student's potentiality for success in school. The three measures (Language, Reasoning, and Quantitative) are combined to obtain a total score, which is then converted to an I.Q. score. The test includes problems in word group, vocabulary, reasoning, letter series, and mathematical problems.

A comprehensive description of the TEA including reliability and validity data can be obtained from the TEA Manual.<sup>71</sup>

---

<sup>69</sup>Frank S. Freeman, Theory and Practices of Psychological Testing (New York: Holt Rinehart and Winston, 1962). p. 274.

<sup>70</sup>Wechsler.

<sup>71</sup>Tests of Educational Ability, Technical Supplement (Chicago: Science Research Associates, Inc., 1964). p. 45.

## California Achievement Tests

The Reading Vocabulary and Reading Comprehension sections of the California Achievement Tests (Total Reading Scores) were used to indicate the students' reading ability. The Reading Vocabulary Test is composed of fifty items, each of which consists of a key word. The following areas are included in the vocabulary: (1) sampling mathematics, (2) science, (3) social science, and (4) general vocabulary items. The Reading Comprehension Section is designed to reveal the following factors: (1) pupil's understanding of what he reads, (2) ability to follow specific instructions, (3) ability to find sources and do reference work, (4) comprehending inferences and drawing valid conclusions from materials read.<sup>72</sup>

The California Achievement Tests are considered adequate as to (1) range of grades covered, (2) aspects and comprehensiveness of subject matter covered, (3) reliability, (4) validity, and (5) standardization by Freeman who considers the tests as one of the sounder batteries available.<sup>73</sup> Anastasi states that the reliabilities of the major tests are adequate for survey purposes.<sup>74</sup>

---

<sup>72</sup>California Achievement Tests Manual (Monterey, California Test Bureau, 1963).

<sup>73</sup>Freeman. p. 499.

<sup>74</sup>Anastasi. p. 470.



### Grade Point Average

A grade point average was computed for each student in the sample. The grades obtained in each academic subject for the first semester (marking period) were averaged and then converted into a number as follows: A = 4, B = 3, C = 2, D = 1, F = 0. Pluses and minuses were disregarded in the computations.

The academic grade point average was used to determine the concurrent validity of the various intelligence tests within each reading ability group.

### Procedure

In December of 1963 the seventh grade students at Stillwater Junior High School, Stillwater, Oklahoma, were divided into "high," "middle," and "low" reading ability groups on the basis of their California Achievement Test Total Reading percentile scores. The "low" group included those students ranking below the 34th percentile. The "middle" group was comprised of students in the 34-66 percentile range. Students with a percentile score above 66 were included in the "high" reading ability group. Forty-two students were randomly selected from the "high" group. There were only 42 students in the entire "low" group and therefore all of those students were used in the study in lieu of a random selection.

Three intelligence tests were administered in counter-balanced order to each subject during the period of January 1,

1964 to March 15, 1964. The fourth test (TEA) was administered by the junior high school staff in November, 1963. Two of the tests administered did not require reading on the part of the testee (Peabody Picture Vocabulary Test and the Wechsler Intelligence Scale for Children). Two of the tests did require reading (TEA) and the Peabody Picture Vocabulary Test (Form B). The Peabody Test (Form B) was altered as to method of presentation so that the testee would be required to read the word to be defined in place of the examiner who would normally pronounce the words. This alteration of the Peabody Test (Form B) was conducted to determine if the necessity of reading the words would handicap the less capable readers (the "low" and possibly the "middle" groups). All of the tests were administered by fully qualified personnel. The results of the Peabody Test (Form A) and WISC were presented to the principal of the junior high school for school use.

An academic grade point average was computed for each student. This grade point average was compiled to assess the degree of predictive efficiency (concurrent validity of the various intelligence tests used in the study).

All of the intelligence tests administered were scored twice to prevent errors.

An analysis of variance was conducted to determine if the scores of each intelligence test exhibited homogeneity of variance within each reading ability group and for the total group. It was necessary to randomly withdraw the scores of six subjects in the "high" group and seven subjects in the "low"

group in order to obtain a condition where the standard deviations did not differ significantly in all four groups. These procedures were necessary to conform to the mathematical requirements of additional formulas to be used (correlations and analysis of variance).

A Pearsonian correlation technique was applied to measure the coefficients of correlation existing between the various intelligence tests within the three reading ability groups and for the total group. This was done to ascertain if the Peabody Test (Form A) would exhibit a significantly greater degree of correlation with the WISC than would the two I.Q. tests requiring reading (TEA and Peabody Test Form (B) within the "low" and "middle" reading ability groups. Also to be determined, was which test would correlate most closely with the WISC for the "high" and total groups.

Pearsonian correlations were applied to evaluate which intelligence test would exhibit the highest degree of concurrent validity with an academic criterion (the grade point average) within each reading ability group and for the total group.

An interrelation matrix between all tests administered and the academic grade point average was derived for each reading ability group and for the total group.

Using Fisher's Z, t-tests were applied to differences in obtained correlations between each test used to ascertain if there were any significant differences existing in the inter-correlation matrix. T-tests were applied to determine if

there were significant differences between correlations obtained between the four intelligence tests and the academic criterion.

T-tests were applied to all correlations in the matrix to ascertain if each correlation differed significantly from zero.

An analysis of variance was conducted to determine if there was a significant difference between the mean scores of the various tests administered within each reading ability group and for the total group.

## CHAPTER IV

### Results

The results of the study are presented according to the following divisions: (1) Variability of the I.Q. Test Scores Within Groups, (2) Correlations Obtained Between the I.Q. Tests, (3) Predictive Validity of the I.Q. Tests, (4) Comparisons of I.Q. Test Intercorrelations, (5) Comparisons of I.Q. Test Intercorrelations with Grade, (6) Comparative analysis of the Differing Levels of Scores Obtained from the I.Q. Tests, (7) Peabody Test, Form A and The Wechsler Test Mean Score Comparisons, (8) Peabody Test Form A and the TEA Mean Score Comparisons, (9) Alternate Forms of the Peabody Test Mean Score Comparisons, and (10) TEA and Peabody Test Form (B) Mean Score Comparisons with the Wechsler Test.

#### Variability of the I.Q. Test Scores Within Groups

An analysis of variance was conducted to determine if there were any significant differences existing between the standard deviations of the I.Q. scores of each intelligence test within each group. The standard deviations derived from the tests within each reading ability group are shown in Table III.

TABLE III

## MEANS AND STANDARD DEVIATIONS OF I.Q.'S

Tests	Reading Groups							
	Low		Middle		High		Total	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
WISC	92.20	11.96	107.55	10.84	116.56	14.37	105.66	15.72
PPVT-A	91.86	9.53	103.86	11.28	114.44	14.37	103.51	14.28
PPVT-B	85.36	12.74	104.10	10.56	115.47	13.86	102.00	17.14
TEA	86.71	9.81	98.12	9.58	112.58	11.37	99.19	14.49

The intragroup comparisons of the Standard Deviations are shown in Table IV.

TABLE IV

## INTRAGROUP COMPARISONS OF STANDARD DEVIATIONS

TABLES OF F - LEVELS<sup>a</sup>

## Reading Groups

Test Comparisons	Low (N=35)	Middle (N=42)	High (N=36)	Total (N=113)
WISC-PPVT-A	1.575	.924	.999	1.125
WISC-PPVT-B	.881	1.054	1.075	.841
WISC-TEA	1.487	1.280	1.600	1.177
PPVT-A-PPVT-B	.560	1.140	1.076	.747
PPVT-A-TEA	.944	1.385	1.599	1.046
PPVT-B-TEA	1.687	1.215	1.486	1.370

<sup>a</sup>None of the values are significant at either the .01 or .05 level.

## Correlations Obtained Between the I.Q. Tests

The intercorrelations between the various I.Q. tests and the academic grade point average within each reading ability group are presented in Tables V, VI, VII and VIII.

TABLE V  
TEST AND ACADEMIC GRADE INTERCORRELATIONS  
FOR THE LOW READING ABILITY GROUP  
(N=35)

Tests	WISC	PPVT-A	PPVT-B	TEA
WISC				
PPVT-A	.73**			
PPVT-B	.45**	.51**		
TEA	.74**	.72**	.46**	
Grade	.48**	.51**	.37*	.66**

\*Values are significant at the .05 level.  
\*\*Values are significant at the .01 level.

TABLE VI

TEST AND ACADEMIC GRADE INTERCORRELATIONS  
FOR THE MIDDLE READING ABILITY GROUP

(N=42)

Tests	WISC	PPVT-A	PPVT-B	TEA
WISC				
PPVT-A	.41**			
PPVT-B	.39*	.49**		
TEA	.54**	.28	.39*	
Grade	.55**	-.07	.05	.46**

\*Values are significant at the .05 level.

\*\*Values are significant at the .01 level.

TABLE VII

TEST AND ACADEMIC GRADE INTERCORRELATIONS  
FOR THE HIGH READING ABILITY GROUP

(N=36)

Tests	WISC	PPVT-A	PPVT-B	TEA
WISC				
PPVT-A	.69**			
PPVT-B	.56**	.86**		
TEA	.67**	.80**	.72**	
Grade	.60**	.57**	.53**	.69**

\*Values are significant at the .05 level.

\*\*Values are significant at the .01 level.



TABLE VIII  
TEST AND ACADEMIC GRADE INTERCORRELATIONS  
FOR ALL READING ABILITY GROUPS COMBINED

(N=113)

Tests	WISC	PPVT-A	PPVT-B	TEA
WISC				
PPVT-A	.75*			
PPVT-B	.70**	.79**		
TEA	.79**	.77**	.75**	
Grade	.71**	.59**	.61**	.78**

\*Values are significant at the .05 level.

\*\*Values are significant at the .01 level.

Within the "low" reading ability group (Table V), all of the correlations between the I.Q. tests were positive and significant at the .01 level. Coefficients of correlation between the WISC and PPVT-A, PPVT-B, and TEA were .73, .45, and .74 respectively.

Within the "middle" reading ability (Table VI), all of the correlations between the I.Q. scores were positive and significant at the .01 or .05 level except the coefficient of correlation between the PPVT-A and TEA ( $r = .28$ ) which was not significant.

All of the Pearson correlations between the I.Q. tests within the "high" reading ability group (Table VII), were significant at the .01 level and positive. The coefficients of correlation between the Wechsler Intelligence Scale for Chil-

dren and PPVT-A, PPVT-B, and TEA were .69, .56, and .67 respectively.

Within the combined reading ability group (Table VIII), all of the intertest correlations were positive and significant at the .01 level. Coefficients of correlation between the WISC and PPVT-A, PPVT-B and TEA were .75, .70 and .79 respectively.

#### Predictive Validity of the I.Q. Tests

Within the "low" reading ability group (Table V), coefficients of correlation of the scores of the WISC, PPVT-A, PPVT-B and the grade point average are .48, .51, .37 and .66 respectively. Correlations between the WISC, PPVT-A and TEA and the grade point average are each significant at the .01 level. The degree of correlation between the grade point average and the PPVT-B is significant at the .05 level.

Within the "middle" reading ability group (Table VI), the WISC and grade point average yield a coefficient of correlation of .55 (positive and significant at the .01 level). The PPVT-B and TEA correlations with the grade point average are .05 and .46 respectively (the second correlation is positive and significant at the .01 level). The correlation between the PPVT-A and grade point average is  $-.07$  and is not significant at the .01 or .05 level of confidence. This negative correlation may be due to the fact that the PPVT-A does not require reading on the part of the testee and reading skills are usually a prerequisite for academic achievement.

Within the "high" reading ability group (Table VII), the following coefficients of correlation were found to exist:

(1) grade point average and WISC,  $r = .60$ ; (2) grade point average and PPVT-A,  $r = .57$ ; grade point average and PPVT-B,  $r = .53$ ; and grade point average and TEA,  $r = .69$ . All of these correlations are positive and significant at the .01 level of confidence.

Within the total group ( $N=113$ ) (Table VIII), the grade point average is found to correlate .71 with the WISC, .59 with the PPVT-A, .61 with the PPVT-B, and .78 with the TEA. All of these  $r$ 's are positive and significant at the .01 level.

#### Comparisons of I.Q. Intercorrelations

After converting the  $r$ 's into Fisher's Z functions, tests of significance were conducted to determine if there were any significant differences existing as to degree of correlation with the WISC among the other tests (PPVT-A, PPVT-B and TEA) within each reading ability group. It was not the purpose of this study to determine if the TEA and PPVT-B differed significantly as to degree of correlation with the WISC and therefore this test was deleted. The data from the statistical comparisons are contained in Tables IX and X.

TABLE IX

PPVT-A and TEA Correlation Comparisons  
 With the WISC Within Each Reading Ability Group<sup>a</sup>

	WISC-PPVT-A		WISC-TEA		CR
	r	z	r	z	
Low Group (N=35)	.73	.91	.74	.96	-.173
Middle Group (N=42)	.41	.44	.54	.60	-.709
High Group (N=36)	.69	.85	.67	.80	.169
All Groups (N=113)	.75	.98	.79	1.1	0.637

<sup>a</sup>None of the critical ratio values are significant at either the .01 or .05 level.

Hypothesis (1) stated that in terms of degree of correlation with the Wechsler Intelligence Scale for Children, there will not be a significant difference at the .05 level of confidence between the Peabody Test (Form A) and the group intelligence test (TEA) scores within the (1a) "low," (1b) "middle," (1c) "high," and (1d) total reading ability groups. The null hypotheses were not refuted.

TABLE X

PPVT-A and PPVT-B Correlation Comparisons  
 With the WISC Within Each Reading Ability Group<sup>a</sup>

	WISC-PPVT-A		WISC-PPVT-B		CR
	r	z	r	z	
Low Group (N=35)	.73	.91	.45	.48	1.74
Middle Group (N=42)	.41	.44	.39	.41	.138
High Group (N=36)	.69	.85	.56	.64	.855
All Groups (N=113)	.75	.98	.70	.87	.796

<sup>a</sup>None of the critical ratio values are significant at either the .01 or .05 level.

Null hypothesis (2) states that in terms of degree of correlation with the Wechsler Intelligence Scale for Children there will not be a significant difference at the .05 level of confidence between the Peabody Picture Vocabulary Test Scores and the written form of the Peabody Test Scores (Peabody Form B altered to require reading) within the (2a) "low," (2b) "middle," (2c) "high" and (2d) combined reading ability groups. These hypotheses are supported by the findings of the t-tests.

## Comparisons of I.Q. Test Correlations with Grade

The data from the statistical comparisons of each I.Q. test with the grade point average are presented in Tables XI, XII, XIII, XIV and IV.

TABLE XI

### WISC and TEA Correlation Comparisons With Grade Within Each Reading Ability Group<sup>a</sup>

	WISC - Grade		TEA - Grade		CR
	r	z	r	z	
Low Group (N=35)	.48	.52	.66	.79	-1.07
Middle Group (N=42)	.55	.61	.46	.50	.515
High Group (N=36)	.60	.69	.69	.85	-.686
All Groups (N=113)	.71	.89	.78	1.05	-1.15

<sup>a</sup>None of the critical ratio values are significant at either the .01 or .05 level.

Hypothesis (3) states that there will be no significant difference at the .05 level of confidence in the coefficients of correlation obtained between the academic grade point average and the Wechsler Intelligence Scale for Children and group intelligence test (TEA) within the (3a) "low," (3b) "middle," (3c) "high," and (3d) total groups. An examination of the critical ratios of Table XI shows no significant differences in the coefficients of correlation with the grade point average were found to exist within any group. Null

hypotheses (3a), (3b), (3c), and (3d) are therefore supported.

TABLE XII

PPVT-A and TEA Correlation Comparisons  
With Grade Within Each Reading Ability Group

	PPVT-A - Grade		TEA - Grade		CR
	r	z	r	z	
Low Group (N=35)	.51	.56	.66	.79	-.924
Middle Group (N=42)	-.07	-.07	.46	.50	-2.52*
High Group (N=36)	.57	.65	.69	.85	-0.843
All Groups (N=113)	.59	.67	.78	1.05	-2.77**

\*Critical ratio value is significant at the .05 level.

\*\*Critical ratio value is significant at the .01 level.

Hypotheses (4) stated that there will be no significant difference at the .05 level of confidence in the coefficients of correlation obtained between the academic grade point average and the PPVT-A and group intelligence test (TEA) within the (3a) "low"; (3b) "middle," (3c) "high" and (3d) total groups. Significant differences as to degree of correlation with grade point average were found to exist between the TEA and PPVT-A within the "middle" and total groups (Table XII). Based on this sampling the TEA would appear to be a more accurate predictor of academic performance than the PPVT-A within these two reading ability groups. Null hypotheses (4b) and (4d) are therefore refuted while null hypotheses

(4a) and (4c) are supported.

TABLE XIII

NISC and PPVT-B Correlation Comparisons  
With Grade Within Each Reading Ability Group

	NISC - Grade		PPVT-B - Grade		CR
	r	z	r	z	
Low Group (N=35)	.48	.52	.37	.39	.549
Middle Group (N=42)	.55	.61	.05	.05	2.48*
High Group (N=36)	.60	.69	.53	.58	.427
All Groups (N=113)	.71	.89	.61	.71	1.31

\*Critical ratio value is significant at the .05 level.

Null hypothesis (5) stated that there will be no significant difference at the .05 level of confidence in the coefficients of correlation obtained between academic grade point average and the Wechsler Intelligence Scale for Children and Peabody Test in Written form (PPVT-B) within the (5a) "low," (5b) "middle," (5c) "high" and (5d) total groups. The findings support null hypotheses (5a), (5c), and (5d). Null hypothesis (5b) is refuted.



TABLE XIV

PPVT-A and PPVT-B Correlation Comparisons  
With Grade Within Each Reading Ability Group<sup>a</sup>

	PPVT-A - Grade		PPVT-B - Grade		CR
	r	z	r	z	
Low Group (N=35)	.51	.56	.37	.39	.698
Middle Group (N=42)	-.07	-.07	.05	.05	-.560
High Group (N=36)	.57	.65	.53	.58	.269
All Groups (N=113)	.59	.67	.61	.71	-0.303

<sup>a</sup>None of the critical ratio values are significant at either the .01 or .05 level.

Null hypothesis (6) stated that there will be no significant difference at the .05 level of confidence in the coefficients of correlation obtained between the academic grade point average and the Peabody Picture Vocabulary Test and Peabody Test in written form (PPVT-B) within the (6a) "low," (6b) "middle," (6c) "high," and (6d) total reading ability groups. No significant difference in the degree of correlation with the grade point average was found within any group and therefore null hypotheses (6a), (6b), (6c), and (6d) are supported.

TABLE XV

WISC and PPVT-A Correlation Comparisons  
With Grade Within Each Reading Ability Group

	WISC - Grade		PPVT-A - Grade		CR
	r	z	r	z	
Low Group (N=35)	.48	.52	.51	.56	.149
Middle Group (N=42)	.55	.61	-.07	-.07	3.04***
High Group (N=36)	.60	.69	.57	.65	.157
All Groups (N=113)	.71	.89	.59	.67	1.61

\*Critical ratio value is significant at the .05 level.

\*\*Critical ratio value is significant at the .01 level.

Within the "middle" group the difference between the WISC and PPVT-A in degree of correlation with the grade point average was found to be significant. The WISC and grade point average was .55 versus a coefficient of correlation of -.07 between the PPVT-A and grade point average. The difference is significant at the .01 level of confidence. The negative correlation between the PPVT-A and grade point average may indicate a lower degree of content validity for the PPVT-A.

Hypothesis (7) states that there will be no significant difference at the .05 level of confidence in the coefficients of correlation obtained between the academic grade point average and the Wechsler Intelligence Scale for Children and PPVT-A within the (7a) "low," (7b) "middle," (7c) "high,"

and (7d) total groups. Null hypotheses (7a), (7c) and (7d) are supported and (7b) is refuted.

Comparative Analysis of the Differing Levels  
of Scores Obtained from the I.Q. Tests

The mean scores of the two I.Q. tests which require reading on the part of the testee (TEA and PPVT-B) were compared with the WISC and PPVT-A as to means obtained. The PPVT-A and the WISC were similarly compared with each other as to mean scores obtained within each group. The results of the first analysis of variance are presented in Table XVI. Significant F-ratios resulted within the "low," "middle," and total groups when the four I.Q. tests were initially compared. Additional F-tests were then conducted to ascertain where these differences existed. Results of these F-tests are presented in Tables XV, XVI, XVII, XVIII, XIX, XX and XXI. The intragroup differences between the mean scores obtained from the various I.Q. tests within each reading ability group are shown in Table XXII.

Peabody Test Form A and the Wechsler Test Mean  
Score Comparisons

When all four I.Q. tests were compared within the "high" group (Table XVI), the resulting F-ratio was .56. The analysis of variance of the scores obtained from the Peabody Picture Vocabulary Test (Form A) and the WISC are presented in Table XVII. The obtained F-ratios for the two tests were

.0176, (within the "low" group), 2.34 (within the "middle"

TABLE XVI

ANALYSIS OF VARIANCE OF MEAN SCORES OBTAINED FROM  
ALL FOUR I.Q. TESTS

Source	df	ss	MS	F
A. WISC, PPVT-A, PPVT-B, and TEA compared within the low reading group				
Between Groups	3	1223.40	407.80	3.32*
Within Groups	136	16729.20	112.00	
B. WISC, PPVT-A, PPVT-B and TEA compared within the middle reading group				
Between Groups	3	1922.90	640.97	5.72*
Within Groups	164	18369.57	112.00	
C. WISC, PPVT-A, PPVT-B, and TEA compared within the high reading group				
Between Groups	3	308.47	102.82	.56
Within Groups	140	25689.50	183.50	
D. WISC, PPVT-A, PPVT-B, and TEA compared within the total group				
Between Groups	3	2505.93	835.31	3.44*
Within Groups	448	108667.07	242.56	

\*Significant at .05 level.

group) and 1.12 (within the total group). None of these values are significant and therefore evidence is presented that the PPVT-A and WISC did not differ as to mean scores obtained within any of the four reading ability groups.

Hypothesis (8) states that there will be no significant difference at the .05 level of confidence between the mean scores obtained from the Peabody Picture Vocabulary Test (Form A) and the mean of the scores obtained from the Wechsler Intelligence Test within the "low" reading ability group, (8b) the "middle" reading ability group, (8c) the "high" reading ability group, and (8d) the total group. All of these null hypotheses are supported by the results of the analysis of variance. (See Table XVII)

TABLE XVII

ANALYSIS OF VARIANCE OF MEAN SCORES OBTAINED FROM  
WECHSLER AND PEABODY FORM A TESTS

Source	df	ss	MS	F
A. WISC and PPVT-A compared within the low reading ability group.				
Between Groups	1	2.05	2.05	.0176
Within Groups	68	7946.88	116.85	
B. WISC and PPVT-A compared within the middle reading ability group.				
Between Groups	1	286.01	286.01	2.34
Within Groups	82	10031.55	122.34	
C. WISC and PPVT-A compared within the high reading ability group.				
Between Groups	1	80.22	80.22	.3886
Within Groups	70	14449.78	206.43	
D. WISC and PPVT-A compared within the total group.				
Between Groups	1	261.28	261.28	1.12
Within Groups	224	52253.44	233.27	

## Peabody Test Form A and the TEA

### Mean Score Comparisons

The mean scores of the PPVT-A and TEA (Table XVIII) were found to be significantly different at the .05 level of confidence within the "low" and "middle" reading ability groups. Within the "low" group the F-ratio was 4.95, for the "middle" group 6.32. No significant difference in means was revealed within the "high" and total groups.

Null hypothesis (9) stated that there will be no significant difference at the .05 level of confidence between the mean scores obtained from the Peabody Picture Vocabulary Test (Form A) and the mean of the scores obtained from the group intelligence Test (TEA) within the (9a) "low," (9b) "middle," (9c) "high," and (9d) total reading ability groups. Null hypotheses (9a) and (9b) are refuted by the findings of the study and null hypotheses (9c) and (9d) are supported.

TABLE XVIII

ANALYSIS OF VARIANCE OF MEAN SCORES OBTAINED FROM  
PEABODY FORM A AND TEA TESTS

Source	df	ss	MS	F
A. PPVT-A and TEA compared within the low reading ability group				
Between Groups	1	462.86	462.86	4.95*
Within Groups	68	6355.43	93.46	
B. PPVT-A and TEA compared within the middle reading ability group				
Between Groups	1	691.44	691.44	6.32*
Within Groups	82	8977.55	109.48	
C. PPVT-A and TEA compared within the high reading ability group				
Between Groups	1	62.35	62.35	.37
Within Groups	70	11747.64	167.82	
D. PPVT-A and TEA compared within the total group				
Between Groups	1	1053.73	1053.73	4.91
Within Groups	224	48095.93	214.71	

\*Significant at .05 level.



TABLE XIX  
ANALYSIS OF VARIANCE OF MEAN SCORES OBTAINED  
FROM PEABODY FORM A AND PEABODY FORM B TESTS

Source	df	ss	MS	F
A. PPVT-A and PPVT-B compared within the low reading ability group				
Between Groups	1	678.91	678.91	5.37*
Within Groups	68	8600.46	126.48	
B. PPVT-A and PPVT-B Compared within the middle reading ability group				
Between Groups	1	1.19	1.19	.01
Within Groups	82	9786.76	119.35	
C. PPVT-A and PPVT-B compared within the high reading ability group				
Between Groups	1	19.01	19.01	.09
Within Groups	70	13945.86	199.23	
D. PPVT-A and PPVT-B compared within the total group				
Between Groups	1	129.38	129.38	.50
Within Groups	224	57494.22	256.67	

\*Significant at .05 level.

The PPVT-A and PPVT-B were found to differ significantly as to mean scores obtained within the "low" reading ability group ( $F = 5.37$ ). No significant difference was found to exist in any of the other three groups (Table XIX).

Null hypothesis (10) stated that there will be no sig-

nificant differences at the .05 level of confidence between the mean scores obtained from the Peabody Picture Vocabulary Test (Form A) and the mean of the scores obtained from the Peabody Test in written form (Form B) within the (10a) "low," (10b) "middle," (10c) "high," and (10d) combined reading ability groups. The findings of the study refute null hypothesis (10a) and support null hypotheses (10b), (10c), and (10d).

#### TEA and Peabody Test Form B Mean Score

#### Comparisons with the Wechsler Test

Table XX contains the F - tests conducted between the WISC and TEA mean test scores.

TABLE XX  
ANALYSIS OF VARIANCE OF MEAN SCORES OBTAINED  
FROM WECHSLER AND TEA TESTS

Source	df	ss	MS	F
A. WISC and TEA compared within the low reading ability group				
Between Groups	1	526.62	526.62	4.41*
Within Groups	68	8128.74	119.54	
B. WISC and TEA compared within the middle reading ability group				
Between Groups	1	1866.86	1866.86	17.84**
Within Groups	82	8582.80	104.67	
C. WISC and TEA compared within the high reading ability group				
Between Groups	1	284.01	284.01	1.69
Within Groups	70	11743.64	167.77	
D. WISC and TEA compared within the total group				
Between Groups	1	2364.43	2364.43	10.35**
Within Groups	224	51172.92	228.45	

\*Significant at .05 level.

\*\*Significant at .01 level.

A significant difference was discovered between the means of the WISC and TEA within the "low" group (F-ratio = 4.41), the "middle" group (F-ratio = 17.84) and the combined group (F-ratio = 10.35). The differences in the "middle" and total groups were significant at the .01 level of confidence. The

difference within the "low" group was significant at the .05 level of confidence.

Null hypothesis (11) stated that there will be no significant difference at the .05 level of confidence between the mean scores obtained from the Wechsler Intelligence Scale for Children and the mean scores obtained from the group intelligence test (TEA) within the (11a) "low," (11b) "middle," (11c) "high," and (11d) total groups. Hypotheses (11a), (11b), and (11d) are refuted while (11c) is supported. (See Table XX).

Table XXI contains the F-test comparisons between the mean scores obtained from the Wechsler and Peabody Form B Tests.

TABLE XXI

ANALYSIS OF VARIANCE OF MEAN SCORES OBTAINED  
FROM WECHSLER AND PEABODY FORM B TESTS

Source	df	ss	MS	F
A. WISC and PPVT-B compared within the low reading ability group				
Between Groups	1	755.71	755.71	4.95*
Within Groups	68	10373.77	152.56	
B. WISC and PPVT-B compared within the middle reading ability group				
Between Groups	1	250.30	250.30	2.19
Within Groups	82	9624.32	114.54	
C. WISC and PPVT-B compared within the high reading ability group				
Between Groups	1	21.13	21.13	.11
Within Groups	70	13941.86	199.17	
D. WISC and PPVT-B compared within the total group				
Between Groups	1	758.39	758.39	2.8046
Within Groups	224	60571.21	270.41	

\*Significant at .05 level.

The F-test conducted for the WISC and PPVT-B I.Q. tests reveals a significant difference between the mean scores of these measures within the "low" reading ability group. The F-ratio reported here was 4.95 and is significant at the .05 level of confidence. No significant findings were discovered

within the other three groups. (See Table XXI).

Null hypothesis (12) stated that there will be no significant difference at the .05 level of confidence between the mean scores obtained from the Wechsler Intelligence Scale for Children and the mean of the scores obtained from the Peabody Test in written form (Form B) within the (12a) "low," (12b) "middle," (12c) "high," and (12d) total groups. Null hypothesis (12b) is refuted and null hypotheses (12b), (12c) and (12d) are supported by the findings. (See Table XXI).

TABLE XXII

INTRAGROUP COMPARISONS OF MEAN I.Q. SCORES

Comparisons	Differences Between Means			
	Low	Middle	High	Total
WISC-PPVT-A	.34	3.69	2.12	2.15
WISC-PPVT-B	6.57*	3.45	1.09	1.42
WISC-TEA	5.49*	9.43**	3.98	6.47**
PPVT-A-PPVT-B	6.23*	.24	1.03	1.51
PPVT-A-TEA	5.15	5.47*	1.86	4.32

\*Significant at .05 level.

\*\*Significant at .01 level.

## CHAPTER V

### DISCUSSION

This chapter presents a discussion of the results and some limitations of the study.

#### Discussion of Results

In terms of degree of correlation with the scores obtained from the Wechsler Intelligence Scale for Children, no significant differences between the I.Q. scores of the PPVT-A, TEA, and PPVT-B were found to exist within any reading ability group. Therefore, in terms of degree of correlation with the WISC, the Peabody Picture Vocabulary Test was not demonstrated to be a superior measure of ability in comparison with the other tests used for readers of differing ability levels.

The PPVT-A, PPVT-B and TEA all demonstrated a positive and significant degree of correlation with the WISC within each reading group and for the total group. Within the total group (N=113), the coefficients of correlation were as follows: (1) WISC and PPVT-A,  $r = .75$ ; (2) WISC and PPVT-B,  $r = .70$ ; and (3) WISC and TEA,  $r = .79$ . In view of the large N in this group, the magnitude of these correlations should indicate considerable overlap as to psychological functions

measured by the different tests. All three I.Q. tests appear to share considerable communality of function with the WISC. The results of the study in regard to the degree of correlation existing between the WISC and PPVT agree closely with the findings of many other validation studies cited in Chapter II, Review of the Literature of this work.

Within the "low" reading ability group (N = 35), the PPVT-A and TEA are almost identical as to degree of correlation with the WISC scores. The PPVT-A and WISC correlation is .73 versus an r of .74 between the TEA and WISC. The degree of correlation found between the PPVT-A and WISC with below average readers in this study agrees closely with the coefficient of correlation found between another picture vocabulary test and the WISC. Smith and Fillmore conducted a study in which Ammons Full Range Picture Vocabulary Test and the WISC were compared with readers who were at least one year below grade level on a reading achievement test. The coefficient of correlation obtained between the WISC and Ammons was .75 (N = 91).<sup>75</sup>

The findings of the study indicate that in regard to predictive validity the WISC coefficient of correlation with school performance (academic grade point average) was significantly higher than was the PPVT-A and grade correlation. The WISC and grade coefficient of correlation was .55 versus

---

<sup>75</sup>Louis Smith and Arlene Fillmore, "The Ammons FRPV Test and the WISC for Remedial Reading Cases," Journal of Consulting Psychology. 1954, 18. p. 123.



an  $r$  of  $-.07$  between the PPVT-A and grade within the "middle" group. This difference was significant at the  $.01$  level of confidence. The TEA also exhibited a significantly greater degree of correlation with grade than did the PPVT-A within the middle group. Within the total group the TEA and grade point average coefficient of correlation was  $.78$ . This  $r$  was significantly greater in magnitude than was the correlation between the PPVT-A and grade ( $r = .59$ ). Within this study the TEA demonstrates the highest degree of predictive validity of any of the four tests utilized.

The superior performance of the group intelligence test (TEA) as to degree of correlation with academic performance is in line with previous research showing a close agreement between academic performance and performance on group I.Q. tests of a primarily verbal composition. For example, the Henman-Nelson has been found to correlate as high as  $.85$  with school grades.<sup>76</sup>

In regard to the comparisons of mean scores obtained from the I.Q. tests, there was no significant difference shown between the mean scores of the PPVT-A and the mean scores of the WISC within any group. The mean scores of these two tests were almost identical within the "low" reading ability group (WISC =  $92.20$ , PPVT-A =  $91.86$ ). Within the "middle" group there was only a difference of  $3.69$  points between the two means (WISC =  $107.55$ , PPVT-A =  $103.86$ ).

---

<sup>76</sup>Freeman, p. 390.

Within the "high" group the difference was 2.12 points (WISC = 116.56, PPVT-A = 114.44). Within the total group the mean score difference between the WISC and PPVT-A was 2.15 (WISC = 105.66, PPVT-A = 103.51).

The I.Q. tests which require reading all yielded mean scores which were significantly different (.05 level) from the WISC within the "low" reading ability group. Within the "low" group the difference between the WISC and PPVT-B mean scores was 6.57 points (WISC = 92.20, PPVT-B = 85.63). The difference between the mean scores of the WISC and TEA within the low group was 5.49 (WISC = 92.20, TEA = 86.71). It is interesting to note that the TEA and PPVT-B (Written Form) both yield scores that are lower than those of the WISC and PPVT-A within the "low" reading ability group. Using the WISC as a standard, it appears that both of the tests requiring reading underestimate the I.Q.'s of below average readers.

The mean scores of the TEA differ significantly from the mean scores of the WISC within the "middle" and total group as well as the "low" group (.01 level). Within the "middle" group the difference between the TEA and WISC mean scores is 9.43 points (WISC = 107.55, TEA = 98.12). Within the total group the TEA and WISC mean score difference is 6.47 points (WISC = 105.66, TEA = 99.19).

The results of this study would appear to validate Neville's contention that the PPVT can serve as an adequate substitute for the administratively more complicated WISC.

Neville found that poor readers in the middle elementary grades tended to make scores on group I.Q. tests requiring reading which were significantly lower than those scores made on individual tests requiring little or no reading (PPVT and WISC).<sup>77</sup> The results of this study which deals with 7th grade students, would appear to substantiate Neville's findings. The findings of this study indicate that the Peabody Picture Vocabulary Test yields scores that are comparable with those of the WISC for students of below average reading ability as well as for those who have no reading handicap. The other tests utilized which required reading on the part of the testee yielded scores which were significantly lower than those of the WISC within the "low" reading ability group; also, the TEA mean scores were significantly different than WISC mean scores within the "middle" and total groups.

#### Limitations

Limitations of the study which should be considered in interpreting the results include the following: (1) The results indicated by the data were obtained from a sample population of 7th grade students of a given institution, at a given period of time, and therefore should not be generalized to populations which are dissimilar. In the absence of further research, the findings of this study should not be

---

<sup>77</sup>Neville, p. 4.

generalized to students of differing age and grade levels.

(2) Limitations were inherent in the instruments used as the validation criteria for the I.Q. test comparisons. However, an attempt was made to choose the most valid and reliable I.Q. test available as the standard to which the other tests would be compared.

## CHAPTER VI

### SUMMARY AND CONCLUSIONS

A summary of the procedure and results of the study are presented in this chapter with the conclusion.

#### Summary

The purpose of the study was to determine if the Peabody Picture Vocabulary Test is a more valid measure of ability than several intelligence tests requiring reading for both students with poorly developed reading skills and those without a reading handicap.

In December of 1963, the seventh grade students of Stillwater Junior High School, Stillwater, Oklahoma, were divided into "high," "middle," and "low" reading ability groups on the basis of their California Achievement Test Total Reading percentile scores. The "low" group included those students ranking below the 34th percentile. The "middle" group was comprised of students in the 34-66 percentile range. Students with a percentile score above 66 were included in the "high" reading ability group. Forty-two students were randomly selected from the "middle" group and a like number were selected from the "high" group. There were only 42 students in the entire "low" group and therefore all

of these students were used in the study in lieu of a random selection.

Three intelligence tests were administered to each subject and a fourth test was administered by the Stillwater Junior High School staff. The tests used were the (1) Peabody Picture Vocabulary Test (Form A), (2) Peabody Picture Vocabulary Test (Form B administered in an altered form which required reading on the part of the testee), (3) Wechsler Intelligence Scale for Children, and (4) Tests of Educational Ability.

An academic grade point average was computed for each student in order to assess the concurrent validity of the various intelligence tests used in the study.

The Wechsler Intelligence Scale for Children (WISC) was used as a standard to which each of the other tests would be compared. Comparative test validity was measured in terms of ability to approximate WISC results within each reading ability group and for the combined group. Validity was also measured in terms of the comparative ability of each test to correlate with school academic performance (the academic grade point average). Comparisons between the WISC and the other I.Q. tests were made in terms of degree of correlation with the WISC scores within each reading ability group and mean I.Q. scores obtained within each reading ability group.

An analysis of variance was conducted to determine if the scores of each intelligence test exhibited homogeneity of variance (homocedasticity) within each reading ability

group and for the total group. It was necessary to randomly withdraw the scores for six subjects in the "high" group and seven subjects in the "low" group in order to obtain a condition where the standard deviations did not differ significantly in all four comparisons. This reduced the total N to 113 ("low" group = 35 subjects, "middle" group = 42 subjects, "high" group = 36 subjects).

A Pearsonian correlation technique was applied to measure the coefficients of correlation existing between the various intelligence tests within the three reading ability groups and for the total group. Pearsonian correlations were applied to evaluate which intelligence test would exhibit the highest degree of correlation with the grade point average. An intercorrelation matrix between all tests administered and the academic grade point average was derived for each reading ability group and for the total group. Using Fisher's Z, t-tests were applied to differences in obtained correlations between each test used to ascertain if there were any significant differences existing in the intercorrelation matrix. T-tests were applied to all correlations in the matrix to ascertain if each correlation differed significantly from zero.

An analysis of variance was conducted to determine if there were any significant differences between the mean scores of the various tests administered within each reading ability group and for the total group.

Hypothesis (1) stated that in terms of degree of cor-

relation with the Wechsler Intelligence Scale for Children, there will not be a significant difference at the .05 level of confidence between the Peabody Test (Form A) and the group intelligence test (TEA) within the (1a) "low," (1b) "middle," (1c) "high," and (1d) total reading ability groups. These null hypotheses were not refuted.

Null hypothesis (2) stated that in terms of degree of correlation with the Wechsler Intelligence Scale for Children there will not be a significant difference at the .05 level of confidence between the Peabody Picture Vocabulary Test scores and the written form of the Peabody Test scores (Peabody Form B) altered to require reading within the (2a) "low," (2b) "middle," (2c) "high," and (2d) combined reading ability groups. These null hypotheses were not refuted.

Null hypothesis (3) stated that there will be no significant difference at the .05 level of confidence in the coefficients of correlation obtained between the academic grade point average and the Wechsler Intelligence Scale for Children and TEA within the (3a) "low," (3b) "middle," (3c) "high," and (3d) total groups. These null hypotheses were all supported.

Hypothesis (4) stated that there will be no significant difference at the .05 level of confidence in the coefficients of correlation obtained between the academic grade point average and the Peabody Picture Vocabulary Test (Form A) and the TEA within the (4a) "low," (4b) "middle," (4c) "high," and (4d) total groups. Null hypotheses (4b) and (4d) were



refuted and (4a) and (4c) supported. The TEA was found to correlate more closely with academic performance than the Peabody Test within the "middle" and combined groups.

Null hypothesis (5) stated that there will be no significant difference at the .05 level of confidence in the coefficients of correlation obtained between the academic grade point average and the Wechsler Intelligence Scale for Children and Peabody Test in written form (PPVT-B) within the (5a) "low," (5b) "middle," (5c) "high," and (5d) total groups. Only null hypothesis (5b) is refuted.

Null hypothesis (6) stated that there will be no significant difference at the .05 level of confidence in the coefficients of correlation obtained between the academic grade point average and the Peabody Picture Vocabulary Test and Peabody Test in written form (PPVT-B) within the (6a) "low," (6b) "middle," (6c) "high," and (6d) total reading ability groups. Hypotheses (6a), (6b), (6c), and (6d) were all supported.

Null hypothesis (7) stated that there will be no significant difference at the .05 level of confidence in the coefficients of correlation obtained between the academic grade point average and the Wechsler Intelligence Scale for Children and Peabody Picture Vocabulary Test (Form A) within the (7a) "low," (7b) "middle," (7c) "high," and (7d) total groups. Hypotheses (7a), (7c) and (7d) were supported.

Hypothesis (7b) was refuted as the WISC exhibited a significantly higher degree of correlation with academic performance than did the Peabody Test.

Null hypothesis (8) stated that there will be no significant difference at the .05 level of confidence between the mean scores obtained from the Peabody Picture Vocabulary Test (Form A) and the mean of the scores obtained from the Wechsler Intelligence Scale for Children within the (8a) "low," (8b) "middle," (8c) "high," and (8d) total groups. All of these null hypotheses were supported.

Null hypothesis (9) stated that there will be no significant difference at the .05 level of confidence between the mean scores obtained from the Peabody Picture Vocabulary Test (Form A) and the mean of the scores obtained from the TEA within the (9a) "low," (9b) "middle," (9c) "high," and (9d) total reading ability groups. Null hypotheses (9a) and (9b) were refuted while (9c) and (9d) were supported.

Null hypothesis (10) stated that there will be no significant difference at the .05 level of confidence between the mean scores obtained from the Peabody Picture Vocabulary Test (Form A) and the mean of the scores obtained from the Peabody Test in written form (PPVT-B) within the (10a) "low," (10b) "middle," (10c) "high," and (10d) combined reading ability groups. Null hypothesis (10a) was refuted while (10b), (10c), and (10d) were supported.

Null hypothesis (11) stated that there will be no significant difference at the .05 level of confidence between

the mean scores obtained from the Wechsler Intelligence Scale for Children and the mean scores obtained from the TEA within the (11a) "low," (11b) "middle," (11c) "high," and (11d) total groups. All four of these hypotheses were refuted by the findings.

Null hypothesis (12) stated that there will be no significant difference at the .05 level of confidence between the mean scores obtained from the Wechsler Intelligence Scale for Children and the mean of the scores obtained from the Peabody Test in written form (Form B) within the (12a) "low," (12b) "middle," (12c) "high," and (12d) total groups. Null hypothesis (12a) was refuted and null hypotheses (12b), (12c), and (12d) were supported.

Limitations of the study which should be considered in interpreting the results are as follows: First, the results indicated by the study were obtained from a sample population of 7th grade students of a given institution at a given period of time. Therefore, generalizations derived from the obtained results should be made with caution and discrimination. Second, one instrument, the Wechsler Intelligence Scale for Children, was used as a primary validation criteria for the I.Q. test Comparisons.

### Conclusions

It was concluded from the findings of this study that in terms of degree of correlation with the Wechsler Intelligence Scale for Children, the Peabody Picture Vocabulary Test was

not demonstrated to be a superior measure of ability in comparison with the other intelligence tests used, within any reading ability group.

In regard to concurrent validity (correlation with academic performance) the WISC and TEA were significantly superior (.05 level) to the Peabody Picture Vocabulary Test within the "middle" reading ability group. Within the total group the TEA demonstrated a significantly higher degree of correlation with academic performance than did the Peabody Picture Vocabulary Test. On an overall basis, the TEA demonstrated the greatest degree of correlation with school academic achievement of any of the intelligence tests utilized in the study.

In regard to mean scores obtained, only the Peabody Picture Vocabulary Test yielded scores that were not significantly different from those obtained on the Wechsler Intelligence Scale for Children within the "low" reading ability group. The TEA yielded scores that were significantly different from the WISC mean scores within the "middle" and total groups as well as the "low" group.

In terms of mean scores obtained, the data presented in this study indicates that the Peabody Picture Vocabulary Test can serve as an adequate substitute for the administratively more complicated WISC with students of below average reading ability as well as with those of average or above average reading proficiency. In this regard, the PPVT (Form A) was demonstrated to be superior to several intelligence

tests which require reading on the part of the testee.

## SELECTED BIBLIOGRAPHY

1. Allen, R. M. and Jones, R. W. "A Suggested Use and Non-Use for the Peabody Picture Vocabulary Test With the Retarded Child." Psychological Reports. XV. (1964) 421-422.
2. Anastasi, Anne. Psychological Testing. New York: The Macmillan Co., 1957.
3. Barbe, Walter, and Griek, Werner. "Correlations Between Reading Factors and I.Q." School and Society. LXXI. (1942) 134-136.
4. Blair, Glenn and Kaman, James. "Do Intelligence Tests Requiring Reading Ability Give Spuriously Low Scores to Poor Readers at the College Freshman Level?" Journal of Educational Research. XXXVI (1942) 280-283.
5. Bleismer, Emery. "A Comparison of Various Capacity Tests Used with Retarded Readers." Elementary School Journal. LVI. (1956) 400-402.
6. Budoff, Milton, and Purseglove, Eleanor. "Peabody Picture Vocabulary Test Performance of Institutionalized Mentally Retarded Adolescents." American Journal of Mental Deficiency. LXVII. 756-760.
7. Burnett, A. "Comparison of the Peabody Picture Vocabulary Test with the Wechsler-Bellevue and Stanford Binet on Educable Mentally Retarded Children and Adolescents." Unpublished Study. Owatonna: Owatonna State School. 1964.
8. California Achievement Tests Manual. Monterey: California Test Bureau. 1963.
9. Corwin, Betty J. "The Influence of Culture and Language on Performance on Individual Ability Tests." Unpublished Study: Northridge: San Fernando Valley State College. 1962.
10. Cronbach, Lee J. Essentials of Psychological Testing. Second Edition. New York: Harper Brothers. 1960.
11. Dunn, Lloyd M. and Brooks, Sayde T. "Peabody Picture

- Vocabulary Test Performance of Educable Mentally Retarded Children." Training School Bulletin, LVII (1960) 35-40.
12. Dunn, Lloyd, and Hottel, John. "Peabody Picture Vocabulary Test Performance on Trainable Mentally Retarded Children." American Journal of Mental Deficiency, LXI. (1961) 448-252.
  13. Durell, D. D. "The Influence of Reading Ability on Intelligence Measures." Journal of Educational Psychology, XXIV. (1933) 412-416.
  14. Freeman, Frank S. Theory and Practice of Psychological Testing. New York: Holt, Rinehart and Winston. 1962.
  15. Garrett, Jane. "Comparisons of the Peabody Picture Vocabulary Test and Wechsler Intelligence Scale for Children." Unpublished Master's Thesis. Nashville: George Peabody College for Teachers. 1959.
  16. Harley, Randall K. and Dunn, Lloyd M. "Comparability of Peabody, Ammons, VanAlstyne and Columbia Test Scores with Cerebral Palsied Children." Exceptional Children, XXVI. (1959) 70-74.
  17. Harris, Albert J. How to Increase Reading Ability. New York: Longmans, Green and Company. 1961.
  18. Hedger, Mable. "An Analysis of Three Picture Vocabulary Tests for Use with the Deaf." Research Monograph. 1964.
  19. Himelstein, Phillip, and Herndon, James. "Comparison of PPVT, WISC, and Academic Achievement Scores Among Educable Mental Defectives." Psychological Reports, VII. (1960) 502.
  20. Kiekligher, R. "Comparison of PPVT and RSB Test Scores of Educable Mentally Retarded Children." Atlanta: State Department of Education. 1964.
  21. Kimbrell, Don. "Comparison of Peabody, WISC, and Academic Achievement Scores Among Educable Mental Defectives." Psychological Reports, VII. (1960) 502.
  22. Klaus, R. and Starke, Christiana. "Experimental Revision of the Peabody Picture Vocabulary Test as a Predictor of First Grade Reading Ability." Unpublished Study. Nashville: Psychology Department, George Peabody College.

23. Lindstrom, A. "A Comparison of the Peabody Picture Vocabulary Test and the Wechsler Intelligence Scale for Children." Studies in Minnesota Education. 1961. 131-132.
24. Manolakes, George, and Sheldon, William. "The Relation Between Reading Test Scores and Language Factors Intelligence Quotients." Elementary School Journal. Feb. (1955) 347-350.
25. Wein, R. "Use of the Peabody Picture Vocabulary Test with Severely Subnormal Patients." American Journal of Mental Deficiency. LXI. (1962) 296-273.
26. Moed, George, Wight, Byron, and James, Patricia. "Inter-test Correlations of the Wechsler Intelligence Scale for Children and Two Picture Vocabulary Tests." Educational and Psychological Measurements. XXII. (1963) 359-363.
27. Moss, James W. "An Evaluation of the Peabody Picture Vocabulary Test with the PMA and 1937 Stanford-Binet on Trainable Children." Unpublished paper. Urbana: University of Illinois Institute for Research on Exceptional Children. 1962.
28. Moss, James. "The Peabody Picture Vocabulary Test with English Children." British Journal of Educational Psychology. XXX. (1960) 82.
29. Mueller, Max. "Effects of Illustration Size on PPVT Test Performance of Visually Limited Children." Exceptional Child. XXIX. (1962) 124-128.
30. Neville, Donald. "The Relationship Between Reading Skills and I.Q. Scores." Unpublished Study. Nashville: George Peabody College for Teachers. 1964.
31. Norris, Raymond C., Hottel, John V., and Brooks, Sayde. "Comparability of Peabody Picture Vocabulary Test Scores Under Group and Individual Administration."
32. Paar, Henry and Lavitt, J. "The Peabody Picture Vocabulary Test. A Study of Congruent Validity." Unpublished Study. Springfield: Springfield College Psychology Department.
33. Pastovic, J. J., and Guthrie, G. M. "Some Evidence on the Validity of the WISC." Journal of Consulting Psychology. XV. (1951) 385-386.
34. Peabody Picture Vocabulary Test Manual. Tentative Edition. Nashville: George Peabody College. 1959.



35. Plattor, Emma, Plattor, Stanton, Sherwood, Clarence, and Sherwood, Sylvia. "Relationship Between Reading Retardation and the Measurement of Intelligence." September (1959) Personnel and Guidance Journal. 49-51.
36. Saslow, H. "The Comparability of the Peabody Picture Vocabulary Test and Revised Stanford-Binet, Form L-M, with Cerebral Palsied Children." Paper read at the American Psychological Association Meeting. New York City. August, 1961.
37. Saslow, H. and Larson, E. "The Comparability of the Peabody and Children's Picture Tests, Stanford-Binet, and Vineland Scales with Cerebral Palsied Children." Paper read at the Rocky Mountain Psychological Association Meeting. 1963.
38. Shaw, James. "Comparability of Peabody Test and WAIS Scores with Schizophrenics without Brain Damage." Unpublished Paper. Nampa: Nampa State School. 1961.
39. Shipe, Dorothy, Cromwell, R. and Dunn, L. "Responses of Emotionally Disturbed and Non-Disturbed Retardates to PPVT Items of Human vs. Non-Human."
40. Smith, Louis, and Fillmore, Arlene. "The Ammons FRPV Test and the WISC for Remedial Reading Cases." Journal of Consulting Psychology. XVIII. (1954).
41. Stake, Robert, and Merens, William. "Reading Retardation and Group Intelligence Test Performance." Exceptional Children. May 1960. 497-501.
42. Strang, Ruth. Problems in the Improvement of Reading in High School and College. New York: McGraw Hill Book Company, 1946.
43. Stroud, James B. "A Note on the Relationship Between Reading and Intelligence Scores," Research and Evaluation in College Reading. Ninth Yearbook National Reading Conference for College and Adults. 1960.
44. Tempero, Howard E. and Ivanoff, John M. "Effectiveness of the Peabody Picture Vocabulary Test with Seventh Grade Pupils." Unpublished paper. University of Nebraska. 1960.
45. Terman, Lewis, and Merrill, Maud. Measuring Intelligence. Boston: Houghton Mifflin. 1937.
46. Tests of Educational Ability Technical Supplement. Chicago: Science Research Association, Inc. 1964.

47. Thorndike, Robert, and Hagan, Elizabeth. Measurement and Evaluation in Psychology and Education. New York: John Wiley and Sons, Inc. 1961.
48. Tobias, Jack and Gorelick, Jack. "The Validity of the Peabody Test as a Measure of Intelligence of Retarded Adults." Training School Bulletin. LVIII. (1961) 92-98.
49. Tyler, Leona E. Tests and Measurements. New Jersey: Prentice-Hall. 1963.
50. Wechsler, David. Wechsler Intelligence Scale for Children Manual. New York: Psychological Corporation. 1949.
51. Weeks, Richard. "Effectiveness of the Peabody Picture Vocabulary Test with College Students." Journal of Educational Research. LVII. (1963)
52. Wheeler, Lester R. "The Relation of Reading to Intelligence." School and Society. LXX. (1949) 225-227.
53. Wheeler, Lester, and Wheeler, Viola. "The Relationship Between Reading Ability and Intelligence Among University Freshmen." Journal of Educational Psychology. 230-237.
54. White, Margaret L. "Eliminating the Fourth Grade Hump in Reading." Monograph on Language Arts. L (1954)
55. Wolfenberger, Wolf. "The Correlation Between PPVT and Achievement Scores Among Retardates; A Further Study." American Journal of Mental Deficiency. LXII (1962) 450-454.

## APPENDIX A

### PEABODY TEST - REGULAR METHOD OF ADMINISTRATION

#### Regular Instructions

#### With Subjects 8 Years of Age and Above

Introduce the test by saying: "I have some pictures to show you." (With mature subjects, say: "I want to find out how large your vocabulary is!") Turn to Example A and say: "See, there are four pictures on this page. Each of them is numbered." (Indicate this by pointing to each in turn.) "I will say a word, then I want you to tell me the number of (or point to) the picture which best tells me the meaning of the word. Let us try one. Tell me the number of (or point to) the picture which best tells the meaning of crib." When a subject makes the desired response, turn to Example B, saying: "That's fine. Now, what number is 'fin'?" Then turn to Example C saying: "Good!" "What number is butterfly?" Then say: "Fine! Now I am going to show you some other pictures. Each time I say a word, you tell the number of (or point to) the picture which best tells the meaning of the word. As we advance through the book you may not be sure you know the meaning of some of the words, but I want you to look carefully at all of the pictures anyway and choose the one you think is right. What number is \_\_\_\_\_?"

## APPENDIX B

### PEABODY TEST - ALTERED METHOD OF ADMINISTRATION

Used for PPVT-B in Written Form

With Subjects 8 Years of Age and Above

Introduce the test by saying: "I have some pictures to show you." (With mature subjects say: "I want to find out how large your vocabulary is.") Turn to Example A and say: "See, there are four pictures on this page. Each of them is numbered." (Indicate this by pointing to each in turn.) Then say: "At the bottom of the page is a word, read this word and tell me the number of (or point to) the picture which best tells the meaning of the word." (The word in this case is crib.) When a subject makes the desired response, turn to Example B saying: "That's fine. Now, which picture best tells the meaning of this word written at the bottom of the page?" Then turn to Example C saying: "Good! Which picture best tells the meaning of this word?" Then say: "Fine! Now I am going to show you some other pictures. Each time you read the word at the bottom of the page and tell me the number of (or point to) the picture which best tells the meaning of the word. As we advance through the book you may not be sure you know the meaning of some of the words you read, but I want you to look carefully at all of

the pictures anyway and choose the one you think is right."

## VITA

Jerry A. Lavitt

**Title:** A Comparative Evaluation of the Peabody Picture Vocabulary Test as a Measure of Ability for Children of Differing Reading Proficiency Levels.

**Major Field:** Student Personnel and Guidance

### Biographical:

**Personal Data:** Born in Rockville, Connecticut, 1931, the son of Paul S. Lavitt and Sarah G. Lavitt.

**Education:** Attended elementary school in Ellington, Connecticut; graduated Classical High School, Springfield, Massachusetts, 1949; received the Bachelor of Arts Degree in Psychology, 1953 from American International College, Springfield, Massachusetts; received the Master of Education Degree in Guidance and Counseling Psychology in 1959, from Springfield College, Springfield, Massachusetts; received the Specialist in Education Degree in Education and Guidance from Springfield College in 1961; attended Oklahoma State University part-time 1961-1963. Resident study at Oklahoma State University 1963-1964.

**Professional Experience:** Worked in family farm and produce business from 1953-1957; served six months active duty in the United States Army as part of the Massachusetts National Guard Training Program, 1957-1958; worked as a guidance counselor in the Westfield, Massachusetts School System from 1959-1963; appointed graduate assistant in the College of Education at Oklahoma State University, 1963. Worked as a guidance counselor at Avon Senior High School 1964-1965. Appointed Director of Pupil Personnel Services of the Avon, Connecticut School System in 1965.

**Professional Organizations:** American Personnel and Guidance Association, Massachusetts Teachers' Association, Connecticut Teachers' Association.