

**A COMPARISON OF THE WECHSLER INTELLIGENCE
SCALE FOR CHILDREN, REVISED AND THE
OTIS-LENNON MENTAL ABILITY TEST,
METROPOLITAN ACHIEVEMENT TESTS AND COURSE GRADES**

BY

ANNABELLE MIDDLETON JONES

A COMPARISON
OF THE WECHSLER INTELLIGENCE SCALE FOR CHILDREN, REVISED
AND THE OTIS-LENNON MENTAL ABILITY TEST,
METROPOLITAN ACHIEVEMENT TESTS AND COURSE GRADES

A Research Paper
Presented to
the Graduate Council of
Austin Peay State University

In Partial Fulfillment
of the Requirement for the Degree
Master of Arts

by
Annabelle Middleton Jones

August, 1978

To the Graduate Council:

I am submitting herewith a Research Paper written by Annabelle Middleton Jones entitled, "A Comparison of the Wechsler Intelligence Scale for Children, Revised and the Otis-Lennon Mental Ability Test, Metropolitan Achievement Tests and Course Grades." I recommend that it be accepted in partial fulfillment of the requirements for the degree of Master of Arts, with a major in Psychology.

Elizabeth H. Stokes
Major Professor

Accepted for the
Graduate Council:

William H. Ellis
Dean of the Graduate School

ACKNOWLEDGEMENTS

The author expresses sincere appreciation to Dr. Elizabeth H. Stokes, Professor of Psychology, Austin Peay State University, who suggested the area of study and provided assistance throughout; to Dr. Garland Blair, Department of Psychology, who developed the computer program which made the statistical research of this project possible.

Appreciation is also extended to the examiners who participated in the original WISC--WISC-R study and whose data has been integrated in the present research: Dr. Stokes, Diana Brent, Bernabe Marrero, Nancy J. Huddleston and Joyce S. Rozier. Additional appreciation goes to Mr. Marrero, who also collected the Otis-Lennon, Metropolitan Achievement and grade data; to Mr. O.B. Moorehead, Coordinator of Instruction and to the students who volunteered to serve as subjects in this study.

Most sincere appreciation goes to my husband and children for their encouragement from the beginning classes to the completion of this research and program of study.

TABLE OF CONTENTS

LIST OF TABLES.....	v
CHAPTER	PAGE
I. INTRODUCTION.....	1
Purpose of the Study.....	3
Hypotheses.....	3
Limitations of the Study.....	4
II. REVIEW OF THE LITERATURE.....	5
III. METHOD.....	14
Subjects.....	14
Description of the Instruments.....	15
Procedure.....	17
IV. ANALYSIS AND INTERPRETATION OF DATA.....	19
V. SUMMARY AND CONCLUSIONS.....	27
Recommendations for Future Research..	30
REFERENCES.....	32

LIST OF TABLES

TABLE		PAGE
1.	Coefficients of Correlations of Scaled Scores and IQs on the WISC-R with Stanines of the Metropolitan Achievement Tests, Intermediate Level, Form G and the Otis-Lennon Mental Ability Test, Elementary II Level, Form K.....	20
2.	Means and Standard Deviations of of the WISC-R, MAT and Otis-Lennon II.....	22
3.	Coefficients of Correlations of Scaled Scores and IQs on the WISC-R with End of Year Course Grades.....	25

CHAPTER I

INTRODUCTION

On February 5, 1975, the California Board of Education imposed a moratorium on the use of intelligence tests in the placement of students in special programs for mentally retarded pupils referred to as the educable mentally retarded. It is the consequence of a lengthy civil rights suit involving the inordinate overrepresentation of minority children in EMR special education programs. In responding to the suit--Larry P. v. Wilson Riles-- a federal district court in San Francisco accepted the allegation that "IQ testing," as then practiced, was prejudicial to black children and tended to place them undeservedly in allegedly stigmatizing programs for the mentally retarded. The Board of Education had little alternative but to extend the moratorium to all children to avert an equal protection test under the Fourteenth Amendment. (Lambert and Meyers, 1977)

Few would adjure the fact that standardized tests are representative of the major group on whom they were standardized. However, the question remains, does the standardized test, though culturally biased, truly predict success or failure in the classroom? Are scores earned on standardized intelligence tests reflective of actual achievement in the classroom? A proper etiological diagnosis of cultural bias may be taken back to the classroom. Perhaps it is there that changes in programs need to be made for the minority population rather than in the testing programs. Thorndike expressed similar sentiment as a defense witness in the Larry P. trial.

If you are going to have a social setting in which the minority and the major bulk of the population are going to function jointly in a common setting, then it seems to me that it is probably inappropriate to provide separate treatment for the minority groupings in the one context (testing program) when you are not going to deal with them as a separate group in another (education). (Condas, 1978)

Sociologist, Jane Mercer, in reference to her support for the inclusion of the Wechsler Intelligence Scale for Children, Revised (WISC-R) in her system of non-discriminatory assessment, stated that it has always been assumed that intelligence test verbal scores correlated well with success in the classroom. A search for supporting research revealed only one published study for elementary age children correlating the Wechsler tests with school grades. Using data which she had in her files, Ms. Mercer found none of the correlations to be very high. The largest correlation is .51. The average correlation between academic grade point average (GPA) and the Verbal Scale of the Wechsler Intelligence Scale for Children (WISC) is .35. The average is .19 with the Performance Scale and .32 with the Full Scale score (Mercer & Lewis, 1977). Since the WISC has been found to be well correlated with the WISC-R, it is expected that the WISC-R would yield similar correlations with academic GPAs. [Stokes, Brent, Huddleston, Rozier & Marrero (1978) in a study with 59 sixth grade students

found correlation coefficients of .92, Verbal Scale; .76, Performance Scale; and .89, Full Scale.] Although she supports the use of the WISC-R in non-discriminatory assessment, she stated that much more research needs to be conducted to determine if the WISC-R actually does predict the success of children in school as determined by school grades (Mercer, 1978).

Purpose of the Study

It is the purpose of this study to determine if a relationship exists for the given population between the WISC-R, the Otis-Lennon Mental Ability Test and two measures of classroom achievement, the Metropolitan Achievement Tests (MAT) and end of year course grades.

Hypotheses

The hypotheses to be studied is as follows:

1. There is a significant correlation between the WISC-R, Metropolitan Achievement Tests and Otis-Lennon Mental Ability Test.
2. The WISC-R Verbal Scale and Full Scale will have the highest correlation with the Otis-Lennon Mental Ability Test, the Metropolitan Achievement Tests and course grades, while the WISC-R Performance Scale will show the lowest correlations with the other instruments

3. There will be a significant correlation between classroom end of year grades and the WISC-R.

Limitations of the Study

1. Since the students who were tested in the study were volunteers, it cannot be assumed that a representative sample of children were included.

2. Since the sample included only one age group, it cannot be assumed that the results will generalize to other age groups.

3. For purposes of another research project (Stokes, Brent, Huddleston, Rozier, and Marrero, 1978) the WISC was administered to the children included in this study. Some minimal practice effect will therefore be included in approximately one-half of the WISC-R scores.

4. There is a question in the writer's mind as to whether course grades are valid representations of classroom achievement. Grades have been known to be given as rewards and punishment for behavior. Also, a letter grade of "B" in a low level class may not be equivalent to a "B" in a higher level class.

CHAPTER II

REVIEW OF THE LITERATURE

A number of studies have been made correlating the Wechsler Intelligence Scale for Children (WISC) with the WISC-R since 1974 when the revised version was published. Research with the WISC-R which has been comparatively neglected involves correlations with academic achievement in the classroom as measured by achievement tests, but in particular, course grades. Since the WISC-R has been normed on a different population than the WISC, new studies need to be made relating the WISC-R to other standards of measurement. This is especially true with the normal child, as the majority of studies have involved children referred for evaluation for special placement.

The present study is concerned with the correlation of the WISC-R, rather than the WISC, and other measures of achievement and ability. Previous studies with the WISC and achievement measures have, in general, shown significant correlations between the WISC scores and academic achievement as measured by standardized tests.

The Wide Range Achievement Test (WRAT) is currently used as a quick individual measure of estimating achievement level when a large battery of tests is being administered

to a given child. With the PPVT, it has been the most popular instrument for correlational studies with the WISC and WISC-R.

Brook's (1977) study indicates that each scale of the WISC predicts achievement in a significant manner as measured by the Wide Range Achievement Test (WRAT). Correlations ranged from .52 to .81.

The WISC scores all correlated at significant levels with all WRAT scales in Hartlage and Steele's (1977) research, ranging from .443 to .751. The study also demonstrated significant correlations between course grades for children enrolled in Grades 1 and 2. Correlations of the WISC Full Scale with arithmetic grades were low, while the correlations with social studies' grades was .794.

Ollendick, Murphy, and Ollendick (1975) found the WISC and WRAT to correlate significantly in a study with 12 to 16 year old male incarcerated delinquents. However, in comparing the WISC to the Peabody Individual Achievement Test (PIAT), the two measures agreed in terms of gross intellectual classification less than 56% of the time. Substitution of one test for the other is questionable.

Stewart and Morris (1977) in a study with 182 emotionally disturbed adolescents (11 to 18 years of age) found

correlations between the WRAT and the WISC to range from .71 to .90. In the same study, WISC and California Achievement Test correlations ranged from .69 to .86, demonstrating significant correlations.

In a study of 84 regularly placed first grade black children (Sewell & Severson, 1975), the Reading and Arithmetic subtests of the Stanford Achievement Test produced average coefficients of .5 and .6, respectively, when compared to the WISC Full Scale scores. The study indicates moderate correlation between the two measures.

A number of studies comparing the WISC and the Peabody Picture Vocabulary Test (PPVT) have been made. In general, these studies have shown a significant correlation between the IQ scores obtained on the two measures of intelligence with correlation higher with the Verbal Scale than with the Performance Scale (Applebaum & Tuma, 1977; Hatch & Covin, 1977; Gage & Naumann, 1965; Moed, Hughes & Lessler, 1965; McArthur & Wakefield, 1968). Lower correlations were found in some studies, particularly with children of low socioeconomic status (Condit, Lewandowski, & Saccuzzo, 1976; Gensemer, Walker, & Cadman, 1976; Pilley, Harris, Miller, & Rice, 1975; Brown & Rice, 1967).

While the Hartlage and Steele (1977) research

found the WISC and course grades to be significantly related, a study by Goldman and Hartig (1976), using data supplied by Jane Mercer, showed low correlations. The study also demonstrated that correlations between the WISC and course grades were higher for white children than black or brown. Correlations were .25, .12 and .14, respectively.

In comparison to the number of studies made with the WISC, the WISC-R studies which have been published are comparatively few. The following studies compare the WISC-R to other standardized measures of achievement and aptitude.

Hartlage and Steele (1977) found the WRAT and WISC-R to be significantly correlated. Reading and arithmetic scores were significantly correlated with the Verbal, Performance and Full Scale IQ scores. Coefficients ranged from .540 to .764. Spelling correlated significantly but at a lower level, .365 to .351.

Brooks' (1977) study demonstrated a significant correlation between all Wechsler scores and reading, spelling and arithmetic achievement as measured by the WRAT. Correlations ranged from .52 to .83.

Covin and Lubimin (1976), in a study of 51 Caucasian children, reported highest correlations to be between

the WISC-R Verbal Scale and WRAT Spelling, Reading and Arithmetic subtest standard scores (.62, .68 and .59, respectively). Lowest correlations were between the WISC-R Performance Scale and WRAT Spelling, Reading and Arithmetic subtests (.29, .35, and .32, respectively).

Schwarting and Schwarting (1977) demonstrated correlations ranging from .61 to .75 in comparing the WRAT subtests with the WISC-R Verbal, Performance and Full Scale scores for children ages 6 to 11. For a second group, ages 12 to 16, the scores correlated equally well except for the Performance score which was somewhat lower for the older group.

In Covin's (1976) study with 48 second grade, low income children considered for placement in classes for the mentally retarded, the PPVT correlated very low and negatively with the WISC-R and was not found to be a useful instrument. No explanation of this unusually low correlation was given.

Covin (1977) in a later study with 50 low achievers found the PPVT to correlate at .59 with the WISC-R Verbal score and as low as .06 with the Performance Scale. Mean IQs were not significantly different but the PPVT tended to be higher than the WISC-R IQs.

Appelbaum and Tuma (1977) found concurrent validity

coefficients between the PPVT and WISC-R for high socioeconomic students of .80 and .83 for Verbal and Full Scale IQs, with a much lower coefficient for Performance IQs, .53. Correlations of the PPVT and the Verbal, Performance and Full Scale WISC-R scores for students of low socioeconomic status were .76, .55 and .74, respectively. Large differences in IQ scores were found between subjects of low and high socioeconomic classes. Therefore, while the PPVT is recommended for obtaining valid IQ estimates, interpretation in relation to socioeconomic class appears to be necessary.

Rozier (1977) in a study with the subjects of the present sample, found the WISC-R Verbal Scale scores correlations with the PPVT, Forms A and B to range from .34 to .76. The Performance Scale correlations ranged from -.01 to .47 with the PPVT. Rozier states that its high correlation with the vocabulary subtest on the WISC-R, which is an established clinical tool, indicates the usefulness of this test.

In a comparison of the Slosson Intelligence Test (SIT) and the WISC-R IQs among 50 special education candidates, Grades 1-4, Covin (1977) found the two measures to be correlated significantly but stated that the SIT overestimated the WISC-R IQs. As in other cited studies,

the WISC-R Verbal Scale was found to be more sensitive to the mental processes measured by the SIT.

Martin and Kidwell (1977) compared the WISC-R, SIT and National Educational Development Test (NEDT) in a study with 33 Caucasian, middle class students. The SIT scores correlated highly with the WISC-R Verbal and Full Scale scores, .82 and .79, respectively, and to a lesser degree with the WISC-R Performance score, .50. The WISC-R Full Scale correlated at .77 with the NEDT.

Of particular interest are studies involving measures correlated in the present research project. Only two studies were found correlating the Otis-Lennon Mental Ability Test and the WISC-R (Covin, 1976) and one study correlating course grades and the WISC-R (Hartlage and Steele, 1977). No published studies were found correlating the WISC-R with the Metropolitan Achievement Tests or any group administered achievement tests which would be popularly used in the public school classroom.

Covin (1976) collected data from 48 2nd grade, 7 year old public school children who were from low income urban families and were being considered for placement in classes for the educable mentally retarded. The Otis-Lennon Mental Ability Test was not found significantly

different from the Wechsler Verbal, Performance and Full Scales. The Otis-Lennon is a non-verbal test at this level.

Correlations of .72, .59 and .76 were computed between the Otis-Lennon I and WISC-R Verbal, Performance and Full Scale IQs, respectively, by Covin (1976) for a group of 119 public school students suspected of mental retardation, grades 1 through 3. The WISC-R Verbal and Full Scales seem to be more sensitive than the Performance Scale to the mental processes measured by the Otis-Lennon. Values for Comprehension, Picture Completion and Performance Scale scores were very low for black children (.18, -.01, .05). For the total sample and for black children, the Otis-Lennon I tended to slightly underestimate the WISC-R Verbal, Performance and Full Scale IQs and to overestimate the WISC-R Performance and Full Scale IQs for white subjects. The author suggests that the data indicates when predicting WISC-R scores from Otis-Lennon I IQs, the examiner must be willing to accept an error of approximately ± 5.5 IQ points, with the expectation that he will be correct within this range about two-thirds of the time.

Hartlage and Steele (1977) found that, in general, the Verbal, Performance and Full Scale scores of the

WISC-R correlated at significant levels with course grades for 36 children who had completed two years of school. The subjects mean age was 7-9 and mean Slosson IQ was 90. They were generally of lower middle socioeconomic class. WISC-R Full Scale IQ did not correlate at a significant level with social studies, and WISC-R Performance IQ did not correlate at a significant level with either reading or social studies grades for first grade. At the second year level, WISC-R verbal scores correlated with all school grades, while the Full Scale scores did not correlate at significant levels with social studies, and the Performance scores did not correlate at significant levels with arithmetic or social studies.

CHAPTER III

METHOD

Subjects

The sample used in this study was composed of students completing sixth grade from public schools in the Clarksville-Montgomery County School System. All participants volunteered to serve as subjects in exchange for football or basketball tickets at Austin Peay State University home games.

Letters were sent home by the children during the last week of the 1976 school year asking for volunteers to participate in the study. Only those children who had signed permission forms from their parents were used as subjects. The parents were also given the opportunity to choose to place or not to place the child's scores in his/her school records (Rozier, 1977).

The original sample tested was composed of sixty-four students of which 30 were female and 34 were male, approximately 75% were white and 25% were black. At the request of the parents of one black child, all parents were assured that no attention would be given to the race of the children in the study. Subjects ranged in age from 11 to 13 years. Students who were included were

in a 19 month age range in order to include children who might have been retained one year for the purpose of securing a more representative sample of sixth grade students.

During March and April, 1976, all sixth grade students in the county including the sample were tested with the Metropolitan Achievement Tests, Intermediate Level, Form G and the Otis-Lennon Mental Ability Test, Elementary II Level, Form K. Data was not collected for the purpose of the present study until the summer of 1978. Some students had moved from the school system and their records were no longer available. Scores earned on the MAT, Otis-Lennon, and end of year course grades were collected for 28 girls and 21 boys of the sample.

Description of the Instruments

The instruments used in the studies with this sample were the WISC; WISC-R; PPVT, Forms A and B; the Metropolitan Achievement Tests, Intermediate Level, Form G; the Otis-Lennon Mental Ability Tests, Elementary II Level, Form K; and end of year course grades. Data from the WISC and PPVT were developed in earlier studies (Stokes, Brent, Marrero, Huddleston and Rozier, 1978; Rozier, 1977).

The Wechsler Intelligence Scale for Children, Revised was published in 1974. It is basically similar

to the WISC. Changes in the revised version were for the purpose of increasing the reliability of the subtests, to eliminate obsolete items or items found to be too difficult, and to include more current language and pictures. The WISC-R is used with children 6 through 16 years of age. The test consists of ten subtests and two alternate subtests from which a Verbal, Performance and Full Scale score can be computed. The average reliability coefficients based on a sample of 2200 are Verbal, .94; Performance, .90; and Full Scale, .96.

The Metropolitan Achievement Tests are designed to evaluate what is being taught in the classroom. Vast amounts of data were collected from leading textbook series, state guidelines and other curriculum sources in developing the topics and questions to be included in the test. Over a three year period, 250,000 students and teachers were involved in developing the test and its normalization. The test topics at the Intermediate Level are reading, language, spelling, mathematics, science and social studies. Five levels have been published. The Intermediate Level is to be used in testing grades 5.0 to 6.9. Reliability coefficients for the ten subtests range from .84 to .96 at the sixth grade level (Durost, et al., 1971).

The Otis-Lennon Mental Abilities Test is constructed

at six levels to measure verbal, numerical and abstract reasoning abilities. "The Otis-Lennon tests are measures of learned or developed abilities in the broadest sense. Performance on the samples of tasks...reflects a complex interaction of genetic and environmental factors..." (Otis and Lennon, 1969). Data amassed over the past 50 years attest to the general popularity of the Otis tests. Approximately 12,000 pupils per grade were used in the standardization of the Otis. The sample was representative of the United States educational system with respect to size and type of system, socioeconomic composition, and geographic region. Reliability coefficients for all grades range from .88 to .95 (Otis and Lennon, 1969).

Procedure

Although the WISC and PPVT data are not included in this study, each child was administered the WISC, WISC-R and Forms A and B of the PPVT. The tests were administered individually by five well trained examiners; three of whom were certified school psychologists and two who were school psychology interns. To control for practice effect approximately one-half of the children were given the WISC and PPVT, Form A first and the other half given the WISC-R and PPVT, Form B first. Each

examiner then retested the same children 30 to 84 days later using the other form of the Wechsler and PPVT.

The Metropolitan Achievement Tests, Intermediate Level, Form G and the Otis-Lennon Mental Ability Test, Elementary II Level, Form K were administered to all sixth grade students in March and April, 1976, in the Clarksville-Montgomery County School System. Data from these tests as well as the end of year course grades earned by the sample of students being considered was collected by the county school psychologist. The results of the research will provide the local county with specific data concerning the functional characteristics of the tests within the Clarksville-Montgomery County school population. In order to protect the privacy of the individual students, the names of the students were not given to the author. The scores were identified only by numbers.

CHAPTER IV

ANALYSIS AND INTERPRETATION OF DATA

A computer was utilized using the Pearson-Product Moment technique to determine the correlation coefficients. Critical values were determined with a one-tailed test. The coefficients were computed from scale scores for the subtests and from the IQ scores for the Verbal, Performance and Full Scale of the WISC-R. The Otis-Lennon scores were computed as IQ scores. The MAT scores were computed as stanine scores. Course grades were translated from letter grades to a number scale of 1 to 13, with 1 being equivalent to "F" and 13 representing the letter grade "A+."

All correlations of the WISC-R, MAT and Otis-Lennon, as shown in Table 1, yielded coefficients at a .005 level of significance. Therefore, the hypothesis of significant correlation is accepted.

The coefficients ranged from .494 (Metropolitan Mathematics Computation subtest correlated with the WISC-R Comprehension subtest) to .872 (comparing the WISC-R Full Scale score with the Metropolitan median score). The critical value for Pearson's correlations at the .005 level for 47 degrees of freedom using a one-tailed test is .365. The lowest correlations were between the WISC-R

TABLE 1

COEFFICIENTS OF CORRELATIONS OF SCALED SCORES AND IQS ON THE WISC-R AND
SUBTEST STANINES OF THE METROPOLITAN ACHIEVEMENT TESTS,
INTERMEDIATE LEVEL, FORM G AND THE OTIS-LENNON ABILITY TEST, ELEMENTARY II LEVEL, FORM K

	Metropolitan Achievement Tests												
	BAT	WD	TL			MATH			MATH	TL	SOC		OTIS-
	MDN	KNOW	READ	READ	LANG	SPELL	CP	CC	PS	MATH	SCI	ST	LENNON
Information	.706	.726	.744	.729	.694	.506	.551	.537	.571	.618	.745	.732	.733
Comprehension	.660	.675	.701	.707	.636	.508	.494	.562	.558	.568	.673	.606	.631
Arithmetic	.734	.526	.594	.590	.729	.626	.758	.723	.743	.768	.647	.640	.601
Similarities	.707	.684	.710	.704	.696	.513	.598	.550	.523	.630	.734	.697	.635
Vocabulary	.749	.800	.777	.789	.705	.620	.556	.569	.554	.598	.815	.725	.753
Verbal IQ	.848	.805	.833	.831	.826	.659	.715	.705	.708	.765	.857	.811	.793
Performance IQ	.753	.687	.727	.706	.721	.534	.609	.628	.633	.671	.598	.671	.716
Full Scale IQ	.872	.811	.847	.837	.845	.655	.726	.729	.732	.783	.794	.807	.819

All Scores $p < .005$

Information and Comprehension subtests and the MAT Spelling and Mathematics subtests, suggesting that different types of information were being assessed. The MAT Reading subtests also correlated with low coefficients when compared with the WISC-R Arithmetic subtests. However, all correlations were above .49.

Covin (1976) found the Otis-Lennon mean IQ to be slightly lower than the WISC-R Verbal, Performance and Full Scale IQs of nonwhite children and to be slightly higher than the Performance and Full Scale IQs for white subjects. The present sample includes both white and nonwhite and no difference was made between the two groups. Covin suggests an error of approximately ± 5.5 points be accepted when predicting WISC-R scores from Otis-Lennon, Level I IQs. The present study found the Otis-Lennon, Level II mean scores to be approximately three points higher than the WISC-R Performance and Full Scale IQs and approximately four points higher than the Verbal IQ. WISC-R and Otis-Lennon correlations ranged from .601 to .819 (see Table 2).

As proposed in the original hypotheses, in all cases the WISC-R Verbal and Full Scale scores correlated more highly than the WISC-R Performance scores with all MAT and Otis-Lennon scores and with course grades.

TABLE 2

MEANS AND STANDARD DEVIATIONS OF THE WISC-R,
METROPOLITAN ACHIEVEMENT TESTS, AND OTIS-LENNON II

	\bar{X}	S.D.
<hr/>		
WISC-R		
Information	10.16	2.71
Comprehension	10.86	2.19
Arithmetic	11.06	3.16
Similarities	10.94	2.71
Vocabulary	10.71	2.50
Verbal IQ	104.08	13.66
Performance IQ	104.86	15.11
Full Scale IQ	104.96	14.69
Metropolitan Achievement Tests, Intermediate Level, Form G		
Word Knowledge	5.22	1.80
Reading	5.86	2.07
Total Reading	5.63	2.02
Language	5.73	1.90
Spelling	6.18	1.94
Math Computation	5.43	2.13
Math Concepts	5.45	2.07
Math Problem Solving	5.59	2.18
Total Math	5.63	2.08
Science	5.78	2.09
Social Studies	5.53	1.85
Battery Median	5.61	1.85
Otis-Lennon Mental Ability Test, Elementary II Level, Form K		
	108.16	14.59
<hr/>		

The WISC-R Verbal, Performance and Full Scale coefficients when compared with the Median MAT score are .848, .753 and .872, respectively. The average correlation between the nine MAT subtests and the Verbal WISC-R score is .769. The average is .645 with the Performance WISC-R score and .772 with the Full Scale score. The Full Scale score correlated better than the Verbal or Performance.

The findings of this study support those of Covin (1976). In his two studies comparing the WISC-R and the Otis-Lennon, the Verbal and Full Scale score were found to be more sensitive than the Performance Scale to the mental processes measured by the Otis-Lennon. Correlations reported by Covin for students in grades 1 to 3 were .72, .59 and .76 for the Verbal, Performance and Full Scale. The present study comparing Level II of the Otis-Lennon and the WISC-R yielded coefficients of .793, .716 and .819, respectively.

In comparing the Wechsler scale with the end of year course grades, the Verbal Scale was the most sensitive measure and the Performance Scale the least sensitive. The mean coefficients for the Verbal, Performance and Full Scale when compared to grades were .503, .370, .477, respectively. Therefore, the hypothesis

that the WISC-R Verbal and Full Scale will have the highest correlation with the MAT, Otis-Lennon and course grades, while the WISC-R Performance Scale will show the lowest correlation with the other instruments and grades is accepted.

The WISC-R Verbal subtests and IQ, Performance and Full Scale IQs correlated significantly at the .05 level, and the majority at the .005 level of significance, with course grades in Reading, Language, Science and Social Studies, ranging from .317 to .665. In Table 3, Spelling and Arithmetic grade correlates were lower, ranging from .208 to .429. Only two correlations were not significant at the .05 level. Arithmetic grades did not correlate significantly with the Similarities and Vocabulary subtests. The WISC-R Arithmetic subtest correlated at .429 with Arithmetic grades. The highest coefficients, .665 and .634, compared the WISC-R Verbal IQ to the Science and Social Studies grades. The other Verbal Scale correlations with grades were: Reading, .553; Language, .495; Spelling, .338; and Arithmetic, .334.

Hartlage and Steele (1977) also found the WISC-R to be significantly correlated with the grades of first and second grade children. The present correlations for sixth grade students are, in general, higher than for the reported first and second grade sample.

TABLE 3

COEFFICIENTS OF CORRELATIONS OF SCALED SCORES AND IQS ON THE WISC-R
WITH END OF YEAR COURSE GRADES

G R A D E S

	Reading	Language	Spelling	Arithmetic	Science	Social Studies
<u>Information</u>	.459**	.389**	.242**	.257*	.615**	.568**
<u>Comprehension</u>	.550**	.485**	.263*	.262*	.597**	.609**
<u>Arithmetic</u>	.483**	.540**	.378**	.429**	.488**	.441**
<u>Similarities</u>	.404**	.334*	.242*	.208	.576**	.502**
<u>Vocabulary</u>	.439**	.317*	.308*	.225	.537**	.583**
<u>Verbal IQ</u>	.553**	.495**	.338*	.334*	.665**	.634**
<u>Performance IQ</u>	.502**	.376**	.239*	.256*	.399**	.447**
<u>Full Scale IQ</u>	.576**	.477**	.312*	.330*	.581**	.587**

*p < .05 (one-tailed test)

**p < .005 (one-tailed test)

The hypothesis that a significant correlation exists between classroom end of year grades and the WISC-R is accepted.

CHAPTER V

SUMMARY AND CONCLUSIONS

The purpose of this study is to determine if a relationship exists for the given population between the WISC-R, Otis-Lennon Mental Abilities Test and two measures of classroom achievement--the Metropolitan Achievement Tests and end of year course grades.

Sixty-four sixth graders from public schools in Clarksville-Montgomery County volunteered for the original study. All subjects were individually administered the WISC-R, WISC, and Forms A and B of the PPVT. The data from the WISC-R is included in the present study. The Metropolitan Achievement Tests and Otis-Lennon were administered to all county sixth grade students a few months prior to the administration of the Wechsler tests. Metropolitan and Otis-Lennon scores earned by the sample group were collected as well as their end of year course grades. Records for 49 of the original sample were available.

The statistical analysis of the data yielded the following conclusions:

1. Correlations of the three standardized tests, WISC-R, Metropolitan Achievement and Otis-Lennon, yielded coefficients at the .005 level of significance.

Correlations of the WISC-R and MAT range from .494 to .872. WISC-R and Otis-Lennon correlations ranged from .601 to .819.

2. In all cases, the WISC-R Verbal and Full Scale scores correlated better than the WISC-R Performance scores with all MAT and Otis-Lennon scores and course grades.

3. The WISC-R Verbal, Performance and Full Scale correlated at the .05 level of significance with the end of year course grades. In addition, all the subtests of the WISC-R Verbal Scale were significant at the .05 level with two exceptions. Correlations of arithmetic grades with the WISC-R Similarities and Vocabulary subtests were not significant. Correlations ranged from .208 to .665. The WISC-R correlates were low when comparing the Verbal Scale to spelling and arithmetic grades, .338 and .334, respectively. Moderate correlations were found for the WISC-R Verbal Scale and science, social studies, reading and language grades, .665, .634, .553, and .495, respectively.

The study supports the premise that the WISC-R does correlate with achievement in the classroom as measured by the MAT and course grades. The value of the WISC-R as a diagnostic instrument for the school

psychologist is complemented by the fact that the WISC-R does provide some indication of level of success in the classroom.

A concern of the writer is the question of whether grades can truly be accepted as valid estimates of classroom achievement. Disregarding the subjective nature of grading a student's written and verbal work in the classroom, there are other factors which complicate the validity of a grading system. Informal survey by the author reveals the fact that a letter grade of "B" in a low level sixth grade class is not earned for equivalent achievement as a "B" in a high level sixth grade class. Teachers also may use grades for reward and punishment of behavior, which complicates the issue of the validity of grades as measures of achievement.

Nadine Lambert, Professor of Education, University of California, discusses similar conclusions:

I think teachers' grades are a very unreliable criterion for measuring school achievement. The reason I say that is that from classroom to classroom teachers have different standards for what constitutes a satisfactory grade or an unsatisfactory grade. Their practices differ in the same school from grade level to grade level. In some of our research, we collected the grades of children in five different school districts. The types of things that children are graded on varies from school to school, and even within a school. So you get a very great degree of variability in what is being graded by a teacher. (Condas, 1978)

Although standardized achievement test scores have been criticized because they may be biased for minority group children, the author questions if more bias would be present in standardized test scores than in teacher grades.

Recommendations For Future Research

1. For valid correlations between classroom achievement and standardized tests, the measure of class achievement must be controlled. It is suggested that a future study be made in which the teacher's evaluations of students' achievement be controlled for the purpose of bringing accuracy and uniformity to the study.

2. It is suggested that a future study be made to compare the correlation of other tests commonly used in the classroom and particularly, the Clarksville-Montgomery County School System, to school grades; i.e., the MAT, Otis-Lennon and graduation qualifying exams.

3. The standardized test is criticized as being culturally biased. The present study attempted to make no evaluation as to possible bias due to an agreement with a subject's parent to give no attention to the

race of the children in the study. A future study to evaluate the possibility of bias should be made.

4. A future study should be made to evaluate whether the correlation of measured IQ and achievement in the classroom of boys differs from that of girls.

REFERENCES

- Anderson, D.E., & Flax, M.L. A comparison of the Peabody Picture Vocabulary Test with the Wechsler Intelligence Scale for Children. Journal of Educational Research, 1968, 62, 114-116.
- Applebaum, A.S., & Tuma, J.M. Social class and test performance: Comparative validity of PPVT with the WISC and WISC-R for two socioeconomic groups. Psychological Reports, 1977, 40, 139-145.
- Brooks, C.R. WISC, WISC-R, S-B L&M, WRAT: Relationships and trends among children ages six to ten referred for psychological evaluation. Psychology in the Schools, 1977, 14, 30-33.
- Brown, L.F., & Rice, J.A. The Peabody Picture Vocabulary Test: Validity for EMRs. American Journal of Mental Deficiency, 1967, 71, 901-903.
- Condas, J. Exhibit C: IQ trial defense experts testify. APA Monitor, 1978, 9, 8-9.
- Condit, J., Lewandowski, D.G., & Saccuzzo, D.P. PPVT, estimation of WISC scores, black versus white delinquents with WISC IQs in 80-89 versus 70-79 versus below 70 range. Psychological Reports, 1976, 38, 359-362.

- Covin, T.M. Comparison of Otis-Lennon Mental Ability Test, Elementary I Level and WISC-R IQs among suspected mental retardates. Psychological Reports, 1976, 38, 403-406.
- Covin, T.M. Comparison of SIT and WISC-R IQs among special education candidates. Psychology in the Schools, 1977, 14, 19-23.
- Covin, T.M. Correlations between the Pintner, Otis-Lennon, Peabody, and Wechsler Intelligence Scale for Children - Revised. Psychological Reports, 1976, 39, 1058.
- Covin, T.M. Relationship of the Peabody and WISC-R IQs of candidates for special education. Psychological Reports, 1977, 40, 189-190.
- Covin, T.M., & Lubimin, A.J. Concurrent validity, WRAT and WISC-R, Caucasian children. Perceptual and Motor Skills, 1976, 43, 573-574.
- Durost, W.N., Bixler, H.H., Wrightstone, J.W., Prescott, G.A., Balow, I.H. Metropolitan Achievement Tests: Teacher's Handbook. New York: Harcourt Brace Jovanovich, 1971.
- Gage, G.E., & Naumann, T.F. Correlation of the Peabody Picture Vocabulary Test and the Wechsler Intelligence Scale for Children. Journal of Educational Research, 1965, 58, 466-468.

- Gensemer, I.B., Walker, J.C., Cadman, T.E. Using the PPVT with children having difficulty learning. Journal of Learning Disabilities, 1976, 9, 179-181.
- Goldman, R.D., & Hartig, L.K. The WISC may not be a valid predictor of school performance for primary-grade minority children. American Journal of Mental Deficiency, 1976, 80, 583-587.
- Hartlage, L.C., & Steele, C.T. WISC and WISC-R correlates of academic achievement. Psychology in the Schools, 1977, 14, 15-18.
- Hatch, G.L., & Covin, T.M. Comparability of WISC and Peabody IQs of young children from three heterogeneous groups. Psychological Reports, 1977, 40, 1345-1346.
- Hughes, R.B., & Lessler, K. A comparison of WISC and Peabody scores of Negro and white rural school children. American Journal of Mental Deficiency, 1965, 69, 877-880.
- Kendall, P.C., & Little, V.L. Correspondence of brief intelligence measures to the Wechsler scales with delinquents. Journal of Consulting and Clinical Psychology, 1977, 45, 660-666.
- Lambert, N., Meyers, C.E., & Opton, Jr., E. From California...two views. APA Monitor, 1978, 9, 4-5.

- Martin, J.D., & Kidwell, J.C. Intercorrelations of the Wechsler Intelligence Scale for Children - Revised, the Slosson Intelligence Test, and the National Educational Developmental Test. Educational and Psychological Measurement, 1977, 37, 1117-1120.
- McArthur, C.R., & Wakefield, H.E. Validation of the PPVT with the Stanford-Binet-LM and the WISC on educable mental retardates. American Journal of Mental Deficiency, 1968, 73, 465-467.
- Mercer, J.R. SOMPA Workshop, Atlanta, 1978.
- Mercer, J.R., & Lewis, J.F. SOMPA: Conceptual and Technical Manual. Riverside: Institute for Pluralistic Assessment Research and Training, 1977.
- Moed, G., Wright, B.W., & James, P. Interest correlations of the Wechsler Intelligence Scale for Children and two picture vocabulary tests. Educational and Psychological Measurement, 1963, 23, 359-363.
- Ollendick, D.G., Murphy, M.J., & Ollendick, T.H. Peabody Individual Achievement Test: Concurrent validity with juvenile delinquents. Psychological Reports, 1975, 37, 935-938.

- Otis, A.S., & Lennon, R.T. Technical Handbook, Otis-Lennon Mental Ability Test. New York: Harcourt, Brace and World, Inc., 1969.
- Pilley, J., Harris, C., Miller, J., & Rice, D. Correlations of WISC and PPVT for socioeconomically deprived black special education classes. Psychological Reports, 1975, 37, 139-144.
- Rozier, J.S. Comparison of scores on the Wechsler Intelligence Scale for Children - Revised and the Peabody Picture Vocabulary Test. Unpublished research paper, 1977.
- Schwarting, F.G., & Schwarting, K.R. The relationship of the WISC-R and WRAT: A study based upon a selected population. Psychology in the Schools, 1977, 14, 431-433.
- Sewell, T.E., & Severson, R.A. Intelligence and achievement in first-grade black children. Journal of Consulting and Clinical Psychology, 1975, 43, 112.
- Stewart, D.W., & Morris, L. Intelligence and academic achievement in a clinical adolescent population. Psychology in the Schools, 1977, 14, 513-518.

Stokes, E.H., Brent, D., Marrero, B., Huddleston, N.J.,

Rozier, J.S. A comparison of WISC and WISC-R scores
of sixth grade students: Implications for validity.

Educational and Psychological Measurements, 1978,
38, 469-473.

Wechsler, D. Manual for the Wechsler Intelligence Scale
for Children - Revised. New York: Psychological
Corporation, 1974.