

**A COMPARISON OF THE WISC AND WISC-R
COMPREHENSION SUBTESTS**

BY

NANCY JO HUDDLESTON

A Comparison Of The WISC And WISC-R
Comprehension Subtests

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

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

by
Nancy Jo Huddleston
July 1977

To the Graduate Council:

I am submitting herewith a Research Paper written by Nancy Jo Huddleston entitled "A Comparison of the WISC and WISC-R Comprehension Subtests." I recommend that it be accepted in partial fulfillment of the requirement for the degree of Master of Arts, with a major in Psychology.

Elizabeth H. Staples
Major Professor

Accepted for the
Graduate Council:

William H. Ellis
Dean of the Graduate School

ACKNOWLEDGEMENTS

The author wishes to express sincere appreciation to Dr. Elizabeth H. Stokes, Professor of Psychology, for her invitation to join the WISC/WISC-R research team and for her infinite patience and guidance during the course of the study and preparation of the manuscript. Appreciation is expressed to Dr. Linda B. Rudolph and Dr. Garland E. Blair, Department of Psychology for their encouragement, criticisms and comments.

The author wishes to thank the other members of the research team, Dr. Elizabeth Stokes, Bernie Marrero, Diana Brent, and Joyce Rozier for their contributions of data to this study.

Special appreciation is extended to my husband, Joe, and to my three daughters for their patience, encouragement, and understanding.

TABLE OF CONTENTS

	Page
LIST OF TABLES.....	iv
CHAPTER	
I. INTRODUCTION	1
II. METHOD	9
Subjects	9
Apparatus	9
Procedure	9
III. RESULTS AND DISCUSSION	12
IV. SUMMARY	20
REFERENCES	21

LIST OF TABLES

Table	Page
1. Means and Standard Deviations for the WISC and WISC-R.....	12
2. Correlations and Single-Tailed <u>t</u> -Tests between the WISC and WISC-R.....	14
3. Number of Examiner Questions Resulting in Improved Raw Scores on WISC-R.....	16

CHAPTER I

INTRODUCTION

Since the passage of Public Law 94-142, the Education for All Handicapped Children Act of 1975, school psychologists have become increasingly aware of the importance of making the most accurate assessment of a child's mental abilities. Therefore, when a widely used individualized intelligence measure is revised, much research is necessary to identify and interpret differences in scores for particular children when they are administered both the older and revised versions of the same instrument.

The Wechsler Intelligence Scale for Children (WISC) became the individual intelligence test of choice for children in the 6 to 13 years age range after its development in 1949 (Hamm, Wheeler, McCallum, Herrin, Hunter, & Catoe, 1976). After the publication of the Wechsler Intelligence Scale for Children-Revised (WISC-R) in 1974, many researchers have observed differences between WISC and WISC-R scores when both instruments had been administered to the same children with the WISC-R scores typically being lower than the WISC scores.

With the recognition of these score differences, it is possible that many parents and teachers may have concluded that the lower scores on the WISC-R may have been

reflecting negatively on children's progress or on the effectiveness of educational programs. A second area of concern is that children who may have scored in the borderline range of intellectual functioning when evaluated with the WISC, will after re-evaluation with the WISC-R, become eligible for classes for the educable mentally retarded (Brooks, 1977; Swerdlik, 1977). Before judgements of these kind can be made, full examination of the restandardization, revisions, additions, and deletions, as well as their effect on scores earned, should be researched.

Wechsler (1974) reported research comparing the mean IQ scores on the WISC-R with the Wechsler Preschool and Primary Scale of Intelligence, the Stanford-Binet (Form L-M, 1972 norms), and the Wechsler Adult Intelligence Scale. He did not report research data on a comparison of the WISC with the WISC-R.

Hamm, Wheeler, McCallum, Herrin, Hunter, and Catoe, (1976) were the first to compare the WISC and the WISC-R when used with a specific population. Children of ages 10 and 13 who had been previously identified as Educable Mentally Retarded (EMR) from a rural southeast Georgia area where selected as subjects. Using a counterbalanced design of administration, the researchers found the WISC-R yielded significantly lower Verbal Scale (VS), Performance Scale (PS), and Full Scale (FS) IQ scores for this particular group. No significant score differences were found

between the age groups. The results emphasize the need for additional criteria in addition to the WISC-R scores before special class placement can be recommended.

Brooks (1977) compared the WISC and the WISC-R when administered to 30 children in the age range of 6 to 10 years who had been referred for psychological evaluation from the public schools. Half the children were administered the WISC first and half were administered the WISC-R first. The WISC FS, VS, and PS mean IQs were significantly higher than the WISC-R FS, VS, and PS mean IQs. The mean scaled score for the WISC-R Comprehension subtest was higher than the mean scaled score for the WISC Comprehension subtest, but he did not indicate if the difference was significant.

Blackman, Gerken, and Snider (1977) researched the effects of ethnic group, sex, age, and ability level on the WISC and WISC-R scores of a stratified random sample of 48 elementary children, using a counterbalanced design. No significant differences were found between the WISC and WISC-R PS and FS scores on the two instruments for their total group, but the WISC VS was significantly higher than the WISC-R VS. This difference was considered to have little practical significance. They found no significant differences related to ethnic groups (whites, blacks, and Chicanos), sex, or age. The below average ability group did obtain higher scores on the WISC for all three scales.

A second study by Blackman, Gerken, and Snider (1977) presented results that conflicted with their prior research. WISC and WISC-R scores of 22 children, 4 blacks, 11 whites, and 7 Chicanos were compared. The data were combined with the scores of the 48 children in the first study for a total sample of 70 children. WISC PS and FS scores were significantly higher than WISC-R scores with no significant difference in the VS scores. The WISC subtest scores were significantly higher on Similarities, Arithmetic, Block Design, and Coding with the WISC-R subtest scores significantly higher on Comprehension.

Schwarting (1976) compared the WISC and WISC-R mean IQ scores of 58 randomly selected children from a school population containing grades one to eight. He used a counterbalanced design of administration to negate the influences of practice and maturation. He found the WISC-R Full Scale IQs, Verbal IQs, and Performance IQs were significantly lower than the corresponding WISC IQ scores. Additionally, it was reported that eight out of ten WISC subtest means were significantly higher than the corresponding WISC-R subtest means, with the Comprehension subtest being the exception. The mean WISC-R Comprehension subtest scale score was significantly higher than the mean WISC Comprehension scaled score, and no significant difference was found for the Vocabulary subtests.

Although studies have shown WISC IQ scores to be

consistently higher than WISC-R IQ scores (Brooks, 1977; Hamm, Wheeler, McCallum, Herrin, Hunter, & Catoe, 1976; Schwarting, 1976), a contradictory study by Covin (1977) employing a counterbalanced design examined the comparability of WISC and WISC-R scores for 30 8- and 9-year-old Caucasian children who had been placed in a private child care institution. All children were from low socioeconomic homes that had been or were being dissolved due to crisis. For this atypical population, no significant differences between WISC and WISC-R IQs were present. Covin (1977) warns that these results should not be generalized to more typical populations.

Swerdlik (1977) feels that children today tend to score higher on the WISC, because they are, on the average, better able to answer IQ type questions than children of 25 years ago. He attributes the change to the improved cultural and educational experiences of children today as compared to those who made up the 1949 standardization sample. Another factor that may account for higher WISC scores is that it takes more raw score points on the WISC-R to earn the same WISC scaled score (Wechsler, 1974).

The Comprehension subtest is one subtest in which the method of administration has been changed to such an extent that there is particular concern about the comparability of scores earned on the two subtests. The Comprehension subtest on the WISC (Wechsler, 1949) consisted of 14

questions which could be scored 2, 1, or 0. General scoring guidelines with limited scoring examples were provided. In order to receive the full credit of 2 points on certain items the student had to spontaneously give two ideas, but they could not be asked to give a second idea in response to the question. Unusual responses could be questioned.

The WISC-R Comprehension subtest (Wechsler, 1974) retained nine items from the 1949 edition, including items which were slightly modified by rewording. Several items were eliminated, including some whose content had been questioned by test users and others that were considered too adult-oriented. Eight new questions were added, primarily to strengthen the test for younger children. According to Sattler (1974) there were eight changes that affected the administration, content, and scoring of the Comprehension subtest. Comprehension is now the ninth subtest instead of the second subtest to be administered. Substantial changes were made in the content. When the first subtest item does not result in a 2-point response, the examiner provides an example of a 2-point response. The number of items has been increased from 14 to 17. The discontinuance rule was changed from three consecutive failures to four consecutive failures. One of the major changes in administration was the specific instructions given to ask for a second idea when the child only gives one idea on items requiring two ideas for full credit.

A comparison of conversion tables of raw scores into scaled scores for chronological ages 12-0 to 12-3 in the WISC and WISC-R manuals indicate that for the WISC-R Comprehension subtest, more raw score points are needed for any given scaled score than would be required for that same score on the WISC. Therefore, it appears that the child who is asked for a second idea would earn a higher score on the WISC-R Comprehension subtest than he would have earned on the WISC when the second answer had to be spontaneously given. Conflicting with the possibility of his earning a higher score when asked for the second idea is the fact that it takes more raw score points on the WISC-R to earn the same standard score as on the WISC.

Although, with the exception of Blackman, Gerken, and Snider (1977) and Covin (1977), researchers have found the WISC FS, VS, and PS IQs to be higher than the corresponding WISC-R IQs (Brooks, 1977; Hamm, Wheeler, McCallum, Herrin, Hunter, & Catoe, 1976; Schwarting, 1976) the mean scaled scores for the WISC-R Comprehension subtest have been higher than for the WISC Comprehension subtest scores (Blackman, Gerken, & Snider, 1977; Brooks, 1977; Schwarting, 1976). Based on previous research, it is therefore hypothesized that:

1. The mean WISC Verbal IQ score will be significantly higher than the mean WISC-R Verbal IQ score.

2. The mean WISC-R Comprehension scaled score will be significantly higher than the mean WISC Comprehension scaled score.

3. The mean raw score total for the five questions common to both the WISC and WISC-R that require a second idea for full credit will be significantly higher for the WISC-R.

CHAPTER II

METHOD

Subjects

The subjects were 59 children whose ages ranged from 11-6 to 12-10 with a mean age of 12-2 at the time of the first testing. All subjects were volunteers who had completed the sixth grade. Sixth-grade students in four elementary schools in the Clarksville-Montgomery County School System were invited to participate in the study. Volunteers were given their choice of football or basketball tickets to Austin Peay State University home games for participating in the study. The sample of 59 students contained 32 males and 27 females; 49 white students, 10 black students.

Apparatus

All children were administered the WISC (Wechsler, 1949) and the WISC-R (Wechsler, 1974). Both instruments are individually administered measures of mental ability with the WISC-R being a revised form of the WISC. Directions for administration and scoring of each instrument were followed as set forth in the respective manuals appropriate to each instrument.

Procedure

The sixth-grade students at each of four elementary

schools were given letters during the last week of school explaining the purpose of the experiment and inviting the students to participate. Attached to the letters were forms for parental permission to participate in the study.

Five examiners, three of whom were certified school psychologists and two who were school psychology interns administered the tests. Four of the five examiners were trained in the administration of the instruments by the same university instructor, who was the fifth examiner. Examiners checked each others protocols for scoring and computational accuracy.

Each child was tested in two separate sessions, which ranged from 30 to 61 days between sessions with a mean difference of 42 days. Each child was administered the two instruments by the same examiner in the two sessions. The children were divided into two groups, Group I ($\underline{n} = 31$) who were administered the WISC first and Group II ($\underline{n} = 28$) who were administered the WISC-R first. Each examiner administered the WISC to the first child tested and the WISC-R to the second child tested. Instruments were alternated with each additional child tested. Since three examiners administered an odd number of tests, three more children were administered the WISC first. The scores from both groups were combined to form the Total Group.

The Verbal Scores, Comprehension subtest scaled scores, and the raw score totals for the five Comprehension

subtest questions common to both the WISC and WISC-R that require a second idea for full credit were analyzed using the t-test for related measures with significance set at the .05 level of confidence. Correlation coefficients were computed using the Pearson Product-Moment Correlation technique. The five questions from the Comprehension subtest analyzed were those regarding wood-brick, criminals, beggars-charity, cotton, and senators-congressmen.

CHAPTER III

RESULTS AND DISCUSSION

WISC and WISC-R mean scores for the Verbal IQ, standard scores for the Comprehension subtest, and raw score totals for the five Comprehension subtest questions common to both instruments were compared. Means and standard deviations for the Total Group, Group I and Group II are listed in Table 1.

Table 1
Means and Standard Deviations
for the WISC and WISC-R

Group	<u>n</u> ^a	<u>M</u>	<u>SD</u>
Total	59		
WISC VIQ		108.19	16.15
WISC-R VIQ		106.14	14.91
WISC SS ^b		10.59	3.21
WISC-R SS ^b		10.92	2.55
WISC Raw Score ^c		4.42	1.75
WISC-R Raw Score ^c		5.66	1.96
I	31		
WISC VIQ		104.87	16.47
WISC-R VIQ		104.10	14.84
WISC SS ^b		10.61	3.18
WISC-R SS ^b		10.65	2.27
WISC Raw Score ^c		4.32	1.54
WISC Raw Score ^c		5.48	1.82

Table 1 (continued)

Group	<u>n</u> ^a	<u>M</u>	<u>SD</u>
II	28		
WISC VIQ		111.86	15.25
WISC-R VIQ		108.39	14.92
WISC SS ^b		10.57	3.30
WISC-R SS ^b		11.21	2.85
WISC Raw Score ^c		4.54	1.99
WISC-R Raw Score ^c		5.86	2.12

^aNumbers indicate number of subjects in each group.

^bScaled score means and standard deviations are for Comprehension subtest only.

^cRaw Score refers to raw score totals for the five questions common to both tests that require two ideas for full credit.

Pearson Product-Moment correlations between the WISC and WISC-R scores indicated that corresponding WISC and WISC-R scores for Verbal IQ, Comprehension Scaled Scores, and raw scores for the five questions common to both instruments were significantly correlated beyond the .01 level of confidence. The t-test for related groups indicated that the WISC mean Verbal IQ scores for the Total Group and Group II were significantly higher than the corresponding WISC-R scores; $t(58) = 2.56$, $p < .01$, $t(27) = 2.76$, $p < .01$. No significant difference was found between the WISC and WISC-R Verbal IQ scores for Group I, $t(30) = .77$, $p > .05$.

Table 2

Correlations and Single-Tailed t -Tests
between the WISC and WISC-R

Group	<u>r</u>	<u>df</u>	<u>t</u>
Total			
VIQ	.92	58	2.563*
SS ^a	.71	58	-1.108
Raw Score ^b	.74	58	-6.989*
I			
VIQ	.94	30	.774
SS ^a	.75	30	-.104
Raw Score ^b	.75	30	-5.320*
II			
VIQ	.90	27	2.761*
SS ^a	.69	27	-1.379
Raw Score ^b	.73	27	-4.600*

Note: All correlations were significant beyond the .01 level of confidence.

^aScaled Score (SS) refers to Comprehension subtest scaled score.

^bRaw Score refers to raw score totals for the five questions common to both tests that require two ideas for full credit.

* $p < .01$

Although WISC-R scaled score means were higher than WISC scaled score means, no significant differences were found between the Comprehension scaled score means for any of the three groups. Total Group scaled score means were 10.59 (S.D. = 3.21) for the WISC and 10.92 (S.D. = 2.55) for the WISC-R, $t(58) = -1.108$, $p > .05$. Group I means

were 10.61 (S.D. = 3.18) for the WISC and 10.65 (S.D. = 2.27) for the WISC-R, $t(30) = -.104$, $p > .05$. Group II means were 10.57 (S.D. = 3.30) for the WISC and 11.21 (S.D. = 2.85) for the WISC-R, $t(27) = -1.379$, $p > .05$.

Raw score means for the five Comprehension subtest questions common to each instrument were significantly higher on the WISC-R for each of the three groups. The Total Group mean WISC raw score total was 4.42 (S.D. = 1.75) and 5.66 (S.D. = 1.96) for the WISC-R, $t(58) = -6.989$, $p < .01$. The Group I mean raw score total was 4.32 (S.D. = 1.54) for the WISC and 5.48 (S.D. = 1.82) for the WISC-R, $t(30) = -5.32$, $p < .01$. Group II mean raw score total was 4.54 (S.D. = 1.99) for the WISC and 5.86 (S.D. = 2.12) for the WISC-R, $t(27) = -4.60$, $p < .01$.

Mandatory questioning for a second idea resulted in approximately a 37% increase in responses and a higher raw score for the five questions common to the Comprehension subtests of each instrument for each of the three groups. The results are shown in Table 3.

Table 3

Number of Examiner Questions Resulting
in Improved Raw Scores on WISC-R

Group	No. Responses Questioned	No. Responses Improved	% Improved
I	101	37	37.0
II	80	30	37.5
Total	181	67	37.0

The results obtained for the Total Group and Group II for the comparison of WISC and WISC-R Verbal IQs are consistent with the majority of previous research and support the first hypothesis which stated that the mean WISC Verbal IQ score would be significantly higher than the mean WISC-R Verbal IQ score. The mean Verbal IQ score for Group I was higher for the WISC, but not significantly higher. However, it does point out conflicting evidence as there were two other reported studies (Blackman, Gerken, & Snider, 1977; Covin, 1977) which failed to find a significant difference in the mean Verbal IQ scores of their subjects.

The second hypothesis which stated that the mean WISC-R Comprehension scaled score would be significantly higher than the mean WISC Comprehension scaled score, must be rejected. Although mean scaled scores for the WISC-R Comprehension subtest were higher for each of the three

groups than for the WISC Comprehension subtest, the mean scores were not significantly different. This finding is not consistent with the results obtained by Blackman, Gerken, and Snider (1977) and Schwarting (1976).

The third hypothesis stated that the mean raw score total for the five questions common to both the WISC and WISC-R that require a second idea for full credit would be significantly higher for the WISC-R. This hypothesis is supported. Mean totals for all three groups were significantly greater for the WISC-R. Wechsler's (1974) addition of requiring the examiner to question for a second idea when only one idea was spontaneously given by the examinee resulted in an average increase in raw score points of approximately 37%. Although the questioning for second ideas has been considered as a possible explanation of significantly higher Comprehension subtest scores, no published literature was located that specifically investigated this explanation. The modifications in wording and the expansion and clarification of scoring examples may also have had an effect on the higher number of points scored, but these questions were not specifically addressed by this study. The higher number of raw score points did not result overall in a higher scaled score. One factor contributing to this phenomena is the requirement of more raw score points to obtain any given scaled score on the WISC-R Comprehension subtest for the age

group involved in the study. For the chronological age group 12-0 to 12-3, the WISC requires 15 raw score points for a scaled score of 10. The same scaled score requires 21-22 raw score points on the WISC-R. If the total possible raw score points are taken into consideration, 54% of the total possible raw score points are required for the scaled score of 10 on the WISC while a minimum of 62% of the total possible raw score points are required for a scaled score of 10 on the WISC-R. This increase in raw score points necessary on the WISC-R for the same scaled score as on the WISC may explain why there was no significant difference in the total Comprehension subtest scaled score even though the scores were significantly higher on equivalent questions when a second idea was requested on the WISC-R. Considering the scorable improvement in examined responses when a second idea was requested, it is possible that responses to the additional four questions on the WISC-R that require two ideas for full credit may also result in higher raw scores when a second idea is requested. Since it was possible to investigate only the five questions common to both instruments, it would not be appropriate to state that the effect of questioning had generalized to the other four subtest items.

Based on the findings in this study, it would appear that more information was being required of children on the WISC-R than was required on the WISC. Therefore, when

a child is administered both instruments this factor should be kept in mind and much caution should be exercised when comparing WISC and WISC-R scores.

Future WISC/WISC-R research should examine the changes in the WISC-R for each of the subtests to determine the effect, if any, on the overall Verbal, Performance, or Full Scale IQ scores. Research of this type would appear to be a necessity, since psychologists today are under increasing pressure to secure the most accurate assessments possible of a child's intellectual functioning. The more that is known about any instrument, the greater will be the accuracy in making judgements concerning intellectual functioning.

CHAPTER IV

SUMMARY

The Verbal IQ scores, Comprehension subtest scores, and raw score totals for five Comprehension subtest items common to both the WISC and WISC-R were examined after both instruments had been administered to 59 children. The children, who had all just completed the sixth grade, were volunteers for a research project conducted by a team of five examiners. Data indicate that the WISC VS IQ scores were significantly higher than WISC-R VS IQ scores for the total sample, but for one group, no significant difference was found. Raw score totals for five Comprehension subtest items common to both instruments were significantly increased when second ideas were requested, but the increase in raw score points did not result in significantly higher scaled scores on the WISC-R Comprehension subtest.

REFERENCES

- Blackman, P. B., Gerken, K. C., Snider, B. C. F. A comparison of the WISC and WISC-R for black, white and Chicano children. Paper presented at the National Association of School Psychologists annual meeting, Cincinnati, March 1977.
- 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 (1), 30-33.
- Covin, T. M. Comparability of WISC and WISC-R scores for 30 8- and 9-year-old institutionalized Caucasian children. Psychological Reports, 1977, 40, 382.
- Hamm, H., Wheeler, J., McCallum, S., Herrin, M., Hunter, D., & Catoe, C. A comparison between the WISC and WISC-R among educably mentally retarded students. Psychology in the Schools, 1976, 13, (1), 4-8.
- Sattler, J. M. Assessment of children's intelligence (Rev. reprint). Philadelphia: W. B. Saunders Co., 1974.
- Schwarting, F. G. A comparison of the WISC and WISC-R. Psychology in the Schools, 1976, 13 (2), 139-141.
- Swerdlik, M. A discussion of WISC/WISC-R (reprinted from Illinois School Psychology Newsletter). Communique.

1977, 5 (8).

Wechsler, D. Manual for the Wechsler Intelligence Scale for Children. New York: The Psychological Corporation, 1949.

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