University of Nebraska at Omaha
DigitalCommons@UNO

Student Work

8-1-1977

WISC, WISC-R Differences: Fact or Artifact

Steven Sherrets
University of Nebraska at Omaha

Follow this and additional works at: https://digitalcommons.unomaha.edu/studentwork
Please take our feedback survey at: https://unomaha.az1.qualtrics.com/jfe/form/SV_8cchtFmpDyGfBLE

Recommended Citation

Sherrets, Steven, "WISC, WISC-R Differences: Fact or Artifact" (1977). *Student Work*. 2406.
https://digitalcommons.unomaha.edu/studentwork/2406

This Thesis is brought to you for free and open access by DigitalCommons@UNO. It has been accepted for inclusion in Student Work by an authorized administrator of DigitalCommons@UNO. For more information, please contact unodigitalcommons@unomaha.edu.

WISC, WISC-R DIFFERENCES; FACT OR ARTIFACT?

Field Project

Presented to the

Department of Psychology

and the

Faculty of the Graduate College

University of Nebraska at Omaha

In Partial Fulfillment

of the Requirements for the Degree

Educational Specialist

BY

Steven Sherrets

August 1977

UMI Number: EP73951

All rights reserved

INFORMATION TO ALL USERS
The quality of this reproduction is dependent upon the quality of the copy submitted.

In the unlikely event that the author did not send a complete manuscript
and there are missing pages, these will be noted. Also, if material had to be removed,
a note will indicate the deletion.

UMI EP73951

Published by ProQuest LLC (2015). Copyright in the Dissertation held by the Author.

Microform Edition © ProQuest LLC.
All rights reserved. This work is protected against
unauthorized copying under Title 17, United States Code

ProQuest LLC.
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106 - 1346

FIELD PROJECT

Accepted for the faculty of The Graduate College of the University of Nebraska at Omaha, in partial fulfillment of the requirement for the degree Educational Specialist.

Graduate Committee Richard L. Wikoff Psychology
Name Department

Kenneth C. Berg - Psychology

John J. [illegible] UNMC

Valerie Cook Psychology

Richard L. Wikoff
Chairman

Date

TABLE OF CONTENTS

INTRODUCTION

The Wechsler Intelligence Scale for Children (WISC) was developed by David Wechsler in 1949 as a downward extension of the adult intelligence test called the Wechsler-Bellevue (Wechsler, 1939). The content for the WISC came largely from the Wechsler-Bellevue Form II although there were some new items added at the easier end to provide a floor for the younger child (Edwards, 1972). As the WISC Manual (Wechsler, 1949) points out, the test was intended for children from ages five through 15 and was standardized separately from the Wechsler-Bellevue Form II.

The WISC quickly became one of the most commonly used tests with children who were having learning or emotional problems (Edwards, 1972. pp. 150). This test has continued in wide use and has been revised and restandardized as the Wechsler Intelligence Scale for Children-Revised (WISC-R: Wechsler, 1974). A comparison of the two scales follows:

Standardization

Wechsler Intelligence Scale for Children. Edwards (1972) points out the WISC was standardized on one hundred boys and one hundred girls at each of the age levels from five through 15. An attempt was made to use children who were at mid-year age. The total sample consisted of 2,200 cases. Although many more children were tested, this final 2,200 cases came closest to the age and other sampling requirements. Only white children were used although the 1950 United States census was used to control for factors such as geographic area, urban-rural distribution and the occupational level of the parents.

Some 85 communities were used for the standardization with the testing being accomplished in the schools. Wechsler does not offer a description on how the children were selected or of either the communi-

ties or the schools involved.

A small group of mentally deficient children were included (N=55). The majority of these children came from three schools, two in the mid-west and one on the east coast. The children selected, had IQ's between 50 and 70 points. The entire process took five years to complete.

Wechsler Intelligence Scale for Children-Revised. The WISC-R Manual (Wechsler, 1974) states that standardization was based on stratification from the 1970 United States Census. Age, sex, race, geographic region, occupation of head of household and urban-rural residence were factors controlled for in the 1974 revision.

The sample contained 100 boys and 100 girls at each of eleven age levels from six and a half through 16½ years for a total of 2,200 cases. Each child was within six weeks of his mid-year. Whites and non-whites (Blacks, American Indians and Orientals) were included at the same rate as were found in the 1970 Census for the age tested. Puerto Ricans and Chicanos were included but were categorized as white or non-white by visible physical characteristics. The sample was limited to "normal" children with no institutionalized mental defectives or children with severe emotional problems.

Testing began in 1971 and concluded in 1973 with some 67 testing centers being used. Some 202 examiners administered the tests in 32 states. The task of locating children was left to the examiners.

Characteristics of the Scales

A complete description of the scales shall not be attempted but a comparison is in order. Both scales attempted to measure global intelligence (Wechsler, 1974 p. 5) as a composite of several traits. Both tests abandoned the use of the mental age concept and substituted the

deviation IQ in its place which compared a child with other children from his own age group.

The WISC and the WISC-R both have 12 subtests (six on the Verbal Scale and six on the Performance Scale) with only ten being mandatory. Each are equally weighted to yield three IQ's, Verbal Performance and the composite of the two, the Full Scale IQ. Much of the WISC-R is quite similar to the WISC in terms of test materials. For a complete description of the changes, the reader is referred to the 1974, WISC-R Manual, pages 10 through 16 (Wechsler, 1974).

There are three notable changes: One being the fact that Performance and Verbal tests are to be alternated rather than being given separately as in the WISC; more demonstration and coaching on the WISC-R; and finally, the number of passed items required at each age level to be considered average has increased.

The WISC-R is intended for children six through 16 years of age, while the WISC was appropriate for children aged five through 15 years.

The content changes on the Verbal tests of the WISC-R, were designed to eliminate items which were too difficult, out of date, highly specific, unimportant or culturally unfair. Digit Span is the only subtest from the Verbal section whose content was not altered. Of the 128 items in this section, only, approximately 30 percent are completely new items.

The Performance subtests were altered even less with over 90 percent of the 200 items having appeared in the previous WISC. Content changes were made to update the materials with several items having been redrawn to depict more minority groups including women and sequences involving children. Object assembly and Coding A and B, did not have

any new items added although the former was redrawn and the latter is now printed in two colors in a separate booklet.

Reliability coefficients, standard-errors and subtest intercorrelations are all a function of age (as well as other variables) and are beyond the scope of the present paper to present. For such information, the reader is referred to the respective manuals. Both scales appear to be acceptably reliable, (test-retest and internal consistency) and have similar standard errors. Slightly lower subtest intercorrelations are achieved on the WISC-R.

Interpretative and Validity Studies

While literally volumes of information exists for the proper interpretation of the WISC, information is just beginning to appear on the WISC-R. With the wide-spread use of the WISC and the vast literature that has developed with it, it is critical to gain as much comparative information on the two forms as possible. This is especially true for children who previously have been tested with the WISC and must now be retested with the WISC-R (such as special education students).

Almost exclusively the research that has appeared on the WISC-R since its introduction has been of a comparative nature with the WISC and the Stanford-Binet (S-B: Terman and Merrill, 1973). Only two studies consider the predictive validity of the WISC and WISC-R. Brooks (1977) found the WISC, the WISC-R and the S-B each predicted global achievement on the Wide Range Achievement Test (Jastak and Jastak, 1965) with children (N=30) from six to 10 years of age. Intelligence quotient were lower with children who were re-evaluated with the WISC-R than on the WISC or the S-B and it was pointed out that this was a point of concern to teachers who felt that IQ's reflect their performance as

teachers.

Hartlage and Steele (1977) attempted a similar study with 36 children from seven to nine years of age. They used the WISC and the WISC-R results and correlated them with achievement scores from the WRAT and school grades at the completion of grades one and two. They found both tests appear to be valid global predictors of achievement. They did, however, find that the WISC subtests were somewhat more highly correlated with specific school grades than were the WISC-R subtests.

In a personal communication between the author and Hartlage (1976), the latter suggested that clinically it was his experience that the WISC-R was not as sensitive an instrument to neurological dysfunction as the WISC and had chosen to exclude it from his neuropsychological battery. No research on the use of the WISC-R in such a battery has appeared to date.

Previous Comparative Studies of the WISC-R

Numerous comparative studies have begun to emerge on the WISC-R with most comparisons being between it and the WISC. The WISC-R Manual reports correlation coefficients between the WISC-R (Full-Scale IQ, Verbal IQ, Performance IQ and Subscale scores) and the S-B (Terman and Merrill, 1972), the Wechsler Adult Intelligence Scale (WAIS: Wechsler, 1955) and the Wechsler Preschool and Primary Scale of Intelligence (WPPSI: Wechsler, 1967) for various ages (See Table 1). While the means are not presented the WISC-R Manual does state that in all cases but one (on the S-B), the IQ's are two to six points lower on the WISC-R.

Lower IQ's on the WISC-R are reported by at least 20 other studies. Although Swerdlik (1977) points out, the majority of these studies have some methodology problems ranging from subject selection to design and analysis errors. Table 2 presents a summary of the comparative studies.

Table 1

Coefficients of Correlation of WISC-R, WAIS and WPPSI, Full Scale IQ's Performance IQ's Verbal IQ's and Stanford-Binet IQ's

	WPPSI[1]			WAIS[2]			S-B[3]
WISC-R Test	Full Scale	Performance	Verbal	Full Scale	Performance	Verbal	
Full Scale IQ	.82	.74	.78	.95	.85	.91	.71
Performance IQ	.78	.80	.63	.79	.83	.68	.60
Verbal IQ	.73	.56	.80	.94	.74	.96	.73

Note: From the WISC-R Manual (1974, pp. 48-52).

1. N = 50
   $\bar{x}$ age = 6 years 0 months
2. N = 40
   $\bar{x}$ age = 16 years 11 months
3. N = 118
   Average correlation from four ages (6-16½)

Table 2

Major Findings of WISC-R Comparison Studies

Researcher	Sample	Procedure	Results*
Brooks (1977)	30, 6-10 year old normal children, 10 females, 20 males, 22 white, 8 non-white	All were administered the WISC-R, S-B WISC and WRAT in one to two days.	WISC-R IQ's were 3-11 pts. lower than those obtained on the S-B or the WISC.
Catron & Catron (Note 1)	30 educably mentally retarded subjects	Counterbalanced Design. WISC-WISC-R	WISC IQ's were an average of 7 points higher than WISC-R
Covin (1977)	50 children, 6-12 years old referred for "special education" evaluation. 31 males, 19 females, Blacks-31, whites-19.	The Slosson Intelligence Test (SIT) was administered along with the WISC-R. SIT was given 2-3 weeks prior to the WISC-R.	The SIT IQ's were higher than the WISC-R by an average of 5 pts.
Davis, C. (Note 2)	22 deaf children aged 9-11 years; urban area	Test-retest interval of not less than 1 yr; WISC always administered first.	Performance IQ: 8.4 pts. Verbal Scale not administered because subjects were deaf.
Davis, E. (1977)	54 "special education" children 21, females, 33 males, 6 to 14 years of age.	The WISC had been administered up to 5 years previous to the WISC-R. A "few had the WISC-R administered first	WISC given first. Full Scale: 2 pts. WISC-R given first. Full Scale: 14 points.

Table 2 Continued.

Researcher	Sample	Procedure	Results*
Doppelt & Kaufman (in press)	WISC & WISC-R Standardization sample	Statistical predication study	Age: 6.5-15.5 Verbal IQ: 1.5 pts. Performance IQ: 6 pts. Full Scale IQ: 4 pts.
Gironda (1977)	20, EMR students $\bar{x}$ IQ =64, 12 to 15 years of age. Majority of the subjects were black	WISC scores were obtained from the students folders and were 6 months to 6 years prior to the administration of the WISC-R	Full Scale: 1 pt higher on the WISC-R Verbal: 2 pts. lower on WISC-R Performance: 1 Pt. higher on the WISC-R
Hamm, et al. (1976)	48 EMR students in rural southeast Georgia	Test-retest interval of 39 days; counter-balanced order	Verbal IQ: 6 pts. Performance IQ: 10 pts. Full Scale IQ: 7.5 pts.
Hartlage & Steele (1977)	36, 7-9 year old children. $\bar{x}$ SIT, IQ =90. 68% male 81% black	WISC, WISC-R, WRAT and subject grades were obtained. The WISC was administered 6 months prior to the WISC-R or WRAT	Full Scale: 2 pts Verbal: 1pt. Performance: 2 pts.
Kaufman & Van Hagen (1977)	80, mentally retarded subjects, 6-16 years of age. 52 males, 28 females. White and non-white	49 subjects had previously been tested with the S-B. All 80 were tested with the WISC-R. The test-retest interval was 2-58 months	S-B IQ's were 7 pts. higher than WISC-R.

Table 2 Continued

Researcher	Sample	Procedure	Results*
Kaufman & Weiner (Note 4)	46 low SES black children aged 7 to 10 years, referred for suspected learning and/or behavioral problems to a New York clinic.	Test-retest interval of 7 weeks; counterbalanced order	Verbal IQ: 7 pts. Performance IQ: 8 pts.
Klinge, et al (1976)	32, adolescent subjects 16 males, 16 females $\bar{x}$ age = 14 years	Counterbalanced, design WISC - WISC-R. Test-retest interval 1 week	Full Scale: 3 pts. Verbal: no difference Performance: 5 pts.
Larrabee & Holroyd (1976)	38 high ability fifth-grade students in a California suburb	Counterbalanced order	Verbal IQ: 9.6 pts. Performance IQ: 8.4 pts. Full Scale IQ: 94 pts.
Reschly & Davis (in press)	48 borderline and educable children aged 719-16.1 in rural Iowa	Test-retest range = 5 to 26 months; mean = 17.3 months; WISC always administered first	Verbal IQ: 7.0 points Performance IQ: .04 pts. Full Scale IQ: 4 points.
Schwarting (1976)	58 randomly selected children aged 6-15 yrs. in a suburban Omaha school.	Test-retest range = 60 - 67 days; counterbalanced order.	Verbal IQ: 4.86 pts. Performance IQ: 8.74 pts. Full Scale IQ: 7.49 pts.

Table 2 Continued

Researcher	Sample	Procedure	Results*
Swerdlik (Note 6)	164 children aged 6 to 15.11 years referred to school psychologists because of concerns about their intellectual ability.	Counterbalanced order with specific test-re-test interval of not less than a week nor more than a month.	Verbal IQ: 4.97 pts. Performance IQ: 5198 pts. Full Scale IQ: 5.47 pts.
Swerdlik & Rice (Note 7)	41 cases of EMR and non-EMR children in Lansing, Michigan area	Test-retest time limit of 4 years. WISC always administered first.	Verbal IQ: 3.80 pts. Performance IQ: 2.74 pts. Full Scale IQ: 3.05 pts. Vocabulary: 1.3. pts.
Zimmerman (Note 8)	86 cases of EH and EMR children in California	No test-retest time limit. WISC always administered first.	EH Students: Verbal IQ: 4.9 pts. Performance IQ: 3 pts. Full Scale IQ: 3 pts. Full Scale IQ: 4 pts. EMR Students: Verbal IQ: 3.3 pts. Performance IQ: 2.2 pts. Full Scale IQ: 2.1 pts.

*In all instances, with the exception of Gironda (1977), the WISC-R scores are lower.

Over half of the studies have been completed on educably mentally retarded (EMR) subjects with the majority of these being comparisons done during mandatory re-evaluations for continued special classroom placements. Davis (1977), Gironda (1977), Kaufman and Van Hagen (1977), Reschly and Davis (in press), Swerdlik and Rice (1975) and Zimmerman (1975) all utilized such a method to examine 329 subjects which included whites and non-whites as well as males and females, ranging in age from six to 16. In almost all cases the WISC had been administered prior to the WISC-R. Test-retest intervals ranged from five years to six months with most of the subjects having different examiners for the two tests.

The greatest differences appeared on the Verbal Scale with all of these studies reporting lower IQ's on the WISC-R ranging from one and a half to seven points. Gironda is the only author which reports higher IQ's on either the Performance or Full Scale IQ's of the WISC-R when compared with the WISC. Other investigators report WISC-R Performance IQ's to range from five hundredths of a point to three points lower and Full Scale IQ's, three to 14 points lower. Davis (1977) reported the greatest discrepancy appeared when the WISC-R was administered prior to the WISC.

Using EMR subjects but employing a counterbalanced design Catron and Catron (1976), Kaufman and Weiner (1976), Hamm, Wheeler, McCallum, Herrin, Hunter, and Catoe (1976) and Swerdlik (1975) compared WISC, WISC-R performance of 288 children ranging in age from six to 15, again including white and non-white, males and females. Test-retest intervals ranged from one week to six months.

These authors report consistently lower IQ's on all of the WISC-R scales. No scale was typically reported lower than the others with all averaging seven points lower, the range being from five to 10 points.

Both EMR students and a counterbalanced design were used by Covin (1977) and Hartlage and Steele (1977) although in the former case the WISC-R was compared with the Slosson Intelligence Test (SIT : Slosson, 1963) on 50 children six to 12 years of age and in the latter case WISC, WISC-R, WRAT, and subject grades were compared on 36 children. Slosson Intelligence Test IQ's were found to be higher by an average of five points. WISC IQ's were higher in the Hartlage study by an average of two points.

Kaufman and Van Hagen (1977) obtained previously administered S-B IQ's for 80 EMR students and compared these with Full Scale WISC-R IQ's. The S-B IQ's were on the average seven points higher.

Normal children numbering 88 (+) including some subjects from the WISC-R standardization sample, were administered both the WISC and WISC-R in a counterbalanced design (Doppelt and Kaufman, in press; Schwarting, 1976). Subjects were varied and ranged in age from six to 15. Test-retest period ranged from one day to seven weeks. In all cases IQ's from the WISC-R were lower than those from the WISC. Differences ranged from three to 11 points with the greatest differences appearing on the Performance Scale.

Two studies utilized unusual subjects in WISC, WISC-R comparisons with Larabee and Holroyd (1976) using 38 gifted students and Davis (1977) using the Performance scale only with 22 deaf children. Davis found the WISC-R to be an average of 8.4 points lower while in the Larabee and Holroyd study Verbal IQ's were 9.6 points lower, Performance 8.4 and Full Scale IQ's 9.4 points lower on the WISC-R.

Finally, using juvenile delinquents and adolescent psychiatric patients; Solway, Fruge, Hays, Gryll, and Cody (1976) and Klinge, Rodgiewicz, and Schwartz (1976) found WISC-R IQ's to be two to four points lower than the WISC counterparts. Klinge, et al. also reports a differential practice effect similar to Davis (1977) in that the greatest score

discrepancies appear where the WISC-R is administered prior to the WISC.

So there appears to be ample evidence that IQ's obtained from the WISC-R will, in general, be lower than those obtained from the S-B, WISC or SIT. Insufficient evidence exists at present to compare the predictive validity of the WISC and WISC-R although Hartlage and Steele (1977) did find a slight preference for the WISC. Both the Davis (1977) and the Klinge et al. (1976) studies raise the question as to what extent results may in part be due to a differential practice effect when the WISC-R is administered prior to the WISC in a counterbalance design. The lack of a statistical analysis for such an order effect would seem to be a serious flaw in the majority of the previous studies. The present study was designed to investigate the possibility of a differential practice effect, as well as to generate comparative data on special education students for the Omaha, Nebraska area.

A pilot study was completed prior to the actual investigation. The pilot consisted of using (N = 14) special education students ($\bar{x}$ age = 13.5) who were in EMR classrooms and all of whom had previously been administered the WISC. As part of a mandatory two year re-evaluation, each subject was administered the WISC-R. The subjects were from a variety of metropolitan schools, with a wide variance in terms of socio-economic background. There were four females and ten males all of whom were white and had Full Scale IQ's ranging from 50 to 86 points on the WISC.

The results of the pilot indicated the two scales were measuring similar abilities with a Full Scale correlation coefficient of .90, Verbal of .86 and Performance of .85. However, the WISC-R IQ's were lower (Performance, 1.43; Verbal, 7.8 points; and Full Scale, 5.0).

The obvious problems with the pilot led into the present study in an attempt to confirm or deny the findings with a more methodologically sound design. Still the author expected to find WISC-R IQ's to be lower.

METHOD

Subjects. Test-retest comparisons were obtained on 30 (males = 21, females = 9) special education students assigned to EMR classrooms from a metropolitan school district. All subjects were white. Two subjects were dropped prior to receiving the second test due to unusual circumstances. They had been involved in a tornado which struck the area the previous afternoon and the examiner felt they were too emotionally upset for administrations of the test to obtain valid scores. This left a total of 28 subjects whose ages ranged from 8.7 to 15.6 years with a mean of 11.8. The mean Full Scale IQ of the subjects were 70.6 from the WISC.

The subjects all attended the same school but were from various classrooms. The school itself was an elementary school (K-6) which was located in a suburb of Omaha, Nebraska. The community itself had 7,460 people but was immediately contiguous with Omaha which has a population of 346,929. Several communities surround Omaha making the majority of the county (Douglas) where they are located a largely urban-suburban area with a population of 418,200 and a density of 1,169.5 people per square mile. Eighty-nine percent of the total labor force was involved in non-agricultural wage and salary employment for a per capita personal income of $7,010.00.

Procedure. The subjects were given both the WISC and WISC-R at two-week intervals. Fifteen of the subjects were given the WISC-R first

---

All Demographic data from the Nebraska Statistical Handbook, 1976-77.

and the WISC second while the order of administration was reversed for the other fifteen. (The two subjects which were dropped were in the latter group which made the N = 13.) Standardization was followed closely for each test. All the administrations were given in the same office, all in the afternoon, by the same female psychometrist.

Written parental permission was obtained prior to data collection. No parents opted to exclude their children. In all cases the children were scheduled for the assessment as part of their mandatory re-evaluation required by the State Department of Education.

RESULTS

A two-factor mixed design: repeated measures on one factor, analysis of Variance (Bruning and Kintz, 1968) was computed for each of the scales. Order of administration, test version (WISC vs. WISC-R) and the interaction between order and version were analyzed. Table 3 presents the means for the three scales with and without the order of administration controlled for.

The analysis of the Full Scale IQ's revealed the WISC-R to be significantly lower than the WISC with a mean differential of 3.43 (F = 12.02, df = 1/26, $P < .01$). Neither the order effect nor the interaction between order and version was significant (F = 1, df = 1/26: F = 3.16, df = 1/26 respectively). The Verbal IQ on the WISC-R was significantly lower than that from the WISC with a mean differential of 4.43 points (F = 33.94, df = 1/26, $P < .001$). The order of administration was not significant (F = 1, df = 1/26), but the intereaction between order and version was (F = 4.25, df = 1/26, $P < .05$).

The same results were obtained on the Performance Scale with order failing to reach significance (F = 1, df 1/26). Version was significant (F = 5.75, df 1/26, $P < .01$). The simple main effects analysis reveals that the interaction between version and order in both the Verbal and Performance tests is accounted for by the fact that when the WISC is administered prior to the WISC-R no significant differences are found with a mean differential of 1.46 on the Full Scale IQ, .08 on Performance and 3.69 on the Verbal IQ. However, when the WISC-R precedes the WISC, all of the IQ's are significantly lower on the WISC-R. The mean Full Scale IQ was 5.07 points lower (F = 12.11, df = 1/26, $P < .01$); the mean Verbal IQ was 5.07 points lower (F = 5.38, df = 1/26, $P < .05$); and the mean Performance IQ was 5.87 points lower (F = 15.76, df = 1/14, $P < .001$).

Table 3

Mean WISC and WISC-R Scores

Overall Means and by Order of Administration

	Overall Means		WISC given first; WISC-R, second		WISC-R given first, WISC, second	
	WISC	WISC-R	WISC	WISC-R	WISC	WISC-R
Full Scale	78.43	75.00	76.85	75.39	79.73	74.67
Verbal	75.32	70.89	74.39	70.69	76.13	71.07
Performance	82.64	82.64	84.00	83.92	87.40	81.53

Table 4 indicates the Pearson Correlation Coefficients are very similar for each of the three conditions. The WISC and WISC-R appear to be measuring similar attributes. All the correlations are statistically significant ($P < .01$).

In an effort to compare the profile of the subtest scores across the two tests a three factor, mixed design: repeated measures on two factors (Bruning and Kintz, 1968) analysis of variance was computed which analyzed order of administration, test version and the ten subtests. The subtest scaled scores served as the dependent variable. Table 5 presents the means of the scaled scores from the subtests with and without the order of administration controlled for.

The analysis revealed that the subscale scores on the WISC-R were significantly lower than those on the WISC ($F = 20.34$, $df = 1/28$, $P < .001$). The main effect for order of administration was not significant ($F = 1$). There was a significant difference for the subtest measures main effect ($F = 109.34$, $df = 9/233$, $P < .001$). Multiple comparisons were computed, using the Newman-Keuls test (Kirk, 1968) to control for an increasing alpha level, in order to investigate the differences among the subtests. Ten stepwise comparisons of the subtests were computed with the critical value (W) ranging from .64 (alpha = .05, W_2) to 1.27 (alpha = .01, W_{10}). The subtests differed from each other largely in the Performance vs. Verbal dimensions. Coding was significantly lower than Object Assembly ($P < .05$) but no other differences were noted among the Performance subtests. Vocabulary, Information and Arithmetic were not significantly different from each other but all were significantly lower than Comprehension, Similarities and all of the Performance Subtests ($P < .01$).

Table 4

Correlations with and without Order of Administration Controlled For

	All subjects	WISC first	WISC-R first
Verbal	.74	.85	.86
Performance	.79	.86	.90
Full Scale	.72	.91	.85

Table 5

Scaled Score Means for Subtests With and Without order of Administration Controlled For

	Overall Means		WISC Given First WISC-R, Second		WISC-R Given First WISC, Second	
	WISC	WISC-R	WISC	WISC-R	WISC	WISC-R
Information	5.54	4.46	5.39	4.31	5.67	4.60
Comprehension	6.54	6.25	7.23	6.46	5.93	6.07
Similarities	7.75	6.14	6.61	6.23	8.73	6.07
Arithmetic	5.43	5.07	5.69	5.00	5.20	5.13
Vocabulary	5.18	4.61	4.69	4.46	5.60	4.73
Picture Completion	8.32	7.82	8.39	8.46	8.27	7.27
Picture Arrangement	7.86	8.21	7.39	7.62	8.13	8.73
Block Design	7.00	6.04	6.69	6.08	7.27	6.00
Object Assembly	8.43	8.68	7.92	8.77	8.87	8.60
Coding	8.32	6.46	8.23	7.31	8.64	5.79

Similarities was significantly lower than Object Assembly ($P < .01$) as well as Picture Arrangement, Picture Completion and Block Design ($P < .05$).

The interaction between order and version was signficant ($F = 4.64$, $df = 1/28$, $P < .05$.) This interaction represents the same results which appeared in the two factor analysis of the Verbal and Performance IQ's. The analysis of the simple main effects reveals that when the WISC is administered prior to the WISC-R no significant differences are found between the scales ($F = 3.05$, $df = 1/28$). When the order is reversed, however, the scores are significantly lower on the WISC-R ($F = 23.43$, $df = 1/28$, $P < .01$).

Neither the interaction between order and the subtests ($F = 1$) nor the three way interaction between order, version and the subtests ($F = 1.67$, $df = 9/233$) was significant. The interaction between version and the subtests was significant ($F = 3.54$, $df = 9/233$, $P < .001$). The analysis of the simple main effects revealed that subjects were significantly more likely to score higher on the Coding ($F = 22.56$, $df = 1/233$, $P < .001$), Similarities ($F = 16.89$, $df = 1/233$, $P < .001$), Block Design ($F = 6.09$, $df = 1/233$, $P < .025$) and Information ($F = 7.51$, $df = 1/233$, $P < .01$) subtests of the WISC.

DISCUSSION

The overall results appear to confirm the findings of previous studies as well as those of the pilot in that the WISC seems to yield higher scores than the WISC-R. It would appear that at least three conclusions can be gleaned from the present study. 1.) Children have likely gotten brighter thus raising the norms required for average performance. This is consistent with the findings from the 1972 restandardization of the S-B. In the author's opinion WISC-R IQ's are still likely to be lower than S-B IQ's, as Kaufman and Van Hagen (1977) found when EMR subjects are used. There may be little discrepancies when normal or gifted children are utilized.

2.) The lack of statistical significance when the WISC is administered first suggests there is a differential practice effect and therefore some previous studies may have inflated results. There is undoubtedly a practice effect on the WISC-R similar to that on the WISC (Reger, 1962; Quereski, 1968) which may well be accentuated by the additional demonstration and coaching allowed on the WISC-R. What we are likely seeing in counterbalanced comparative studies are not only lower WISC-R scores but inflated WISC scores which make the tests appear to be more disimilar than they really are. Such a conclusion is supported by the Klinge, et al. (1976) study as well as Davis (1977). Further researchers should be aware of this artifact of design that results in the order by version interaction. Appropriate statistical controls should be utilized as well as an avoidance of short test-retest intervals.

The significantly higher scores on the WISC Coding, Information, Similarities and Block Design subtests are more difficult to explain than

the overall differences between the two tests. Similarities underwent the greatest revision of any of the subtests with nearly half of its 17 items being entirely new or substantially revised on the WISC-R. Information was also revised considerably with over thirty per cent of the items being new or substantially changed from the WISC. While Block Design has only two new items (out of 11 total) the time allowed on the designs requiring nine blocks has been reduced by 30 seconds thus making it more difficult.

The difference in the Coding scores is the most puzzling of all to explain. The content of the test and the instructions remain identical to those in the WISC. The only change has been the introduction of color and the use of a separate sheet. The introduction of color may have a slowing effect on the subject's performance.

The differences on the subtests are likely the result of normative changes, content changes, practice effects and even possibly type-one errors. The scaled scores have changed with the new standardization sample which, while it includes the minority groups excluded from the WISC, has excluded mentally defectives. Such an exclusion may in part have influenced the results of comparative studies which use mentally retarded subjects.

The third major conclusion has little to do with the two tests but is concerned with the subjects themselves. It seems to be the case that for at least these EMR subjects any verbal, culture bound tasks will prove to be more difficult than non-verbal ones. This should be taken into account, particularly with EMR subjects, when comparing the WISC-R with other measures of intelligence or achievement. When it is compared with a test that is largely measuring non-verbal ability WISC-R scores will

appear lower than expected from the Full Scale IQ.

To suggest the WISC-R yields lower scores is not to say it is not as valid as its predecessor. Surely research into the predictive validity of the tests will be required to make such determinations. It is reasonable to assume that, as the 1972 Stanford-Binet norms have indicated, that children know more today than they did in 1949. Still the lower scores on the WISC-R will result in more children being placed in classrooms for the retarded and possibly less movement out of such classrooms for mainstreaming. This will be particularly true for children that fall at the borderline level required for such classification. Evaluation of teaching programs and placement decisions may have to be made on more than IQ alone, as well they should be.

Certainly more research is called for which hopefully will be of better quality than that currently appearing. Factor analytic studies begun by Kaufman (1975) should be continued as well as many, many studies of clinical and predictive validity. It will take years before we know as much about the WISC-R as we do about the WISC. In the meantime, caution should be exercised in assuming comparability of scales or previous research on the WISC.

REFERENCES

REFERENCES

Barclay, A., Carolan, P.A. A comparative study of the Wechsler Intelligence Scale for Children and the Stanford-Binet Intelligence Scale Form L-M. Journal of Consulting Psychology, 1966, 30, 563.

Brooks, Clarence. 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, Jan. 1977, Vol. 14, No. 1.

Bruning, James, and Kintz, B. Computational Handbook of Statistics. Glenview, Illinois, Scott, Foresman and Co. 1968.

Covin, Theron. Comparison of SIT and WISC-R IQs among special education candidates. Psychology in the Schools, Jan. 77, Vol. 14, No. 1.

Davis, Everett. Matched pair comparison of WISC and WISC-R scores. Psychology in the Schools, April, 77. Vol. 14, No. 2.

Doppelt, J.E., and Kaufman, A.S. Estimation of the differences between WISC-R and WISC IQs. Journal of Educational and Psychological Measurement, in press.

Edwards, Allen, Individual Mental Testing: Part II Measurement, San Francisco, Intext Educational Publishers, 1972.

Gironda, Ronald. A comparison of WISC and WISC-R results of urban educable mentally retarded students. Psychology in the Schools, July, 1977. Vol., 14, No. 3.

Hamm, H., Wheeler, J., McCallum, S., Herrin, M., Hunter, D., and Catoe, c., A comparison between the WISC and WISC-R among educable mentally retarded students. Psychology in the Schools. 1976, Vol., 13, 4-8.

Hartlage, Lawrence and Steele, Carol. WISC and WISC-R correlates of academic achievement. Psychology in the Schools, Jan., 77. Vol., 14, No. 1.

Jastak, J.F., and Jastak, S.R. The Wide Range Achievement Test, Wilmington: Guidance Associates, 1965.

Kaufman, Alan, and Van Hagen, John. INvestigation of the WISC-R for use with retarded children: correlation with the 1972 Stanford-Binet and comparison of WISC and WISC-R profiles. Psychology in the Schools. Jan. 77. Vol. 14, No. 1.

Kirk, Roger, Experimental Design; Procedures for the Behavioral Sciences. Belmont, California, Brooks/Cole Co., 1968.

Klinge, V., Rodgiewicz, T., and Schwartz, L. Comparison of the WISC and WISC-R on a psychiatric adolescent inpatient sample. Journal of Abnormal Child Psychology, Vol. IV, No. 1, 1976.

Krichev, A. A revision that really is--The WISC-R. Psychology in the Schools, 1975, 12, 126-28.

Larrabee, G. and Holroyd, R. Comparison of WISC and WISC-R using a sample of highly intelligent children. Psychological Reports, 1976, 38, 1077-80.

Overeshi, Mohammed. Practice effects on the WISC subtest scores and IQ estimates. Journal of Clinical Psychology, Vol., 24, No. 1, 1968, 79-85.

Reger, R. Repeated measurements with the WISC. Psychological Reports, Vol. 11, No. 2, 1962, 418.

Reschly, D. and Davis, R. Comparability of WISC and WISC-R scores. Journal of Clinical Psychology, in press.

Schwarting, F. A comparison of the WISC and WISC-R Psychology in the Schools, 1976, 13, 139-41.

Slosson, R. Slosson Intelligence Test (SIT) for Children and Adults. East Aurora, N.Y.: Slosson Educational Publishers, 1963.

Swerdlik, Mark. The question of the comparability of the WISC and WISC-R: Review of the research and implications for school psychologists. Psychology in the Schools, July, 1977, Vol. 14, No. 3.

Terman, L. and Merrill, M. Stanford-Binet Intelligence Scale. Boston: Houghton Mifflin, 1973.

Wechsler, D. The Measurement of Adult Intelligence. Baltimore: Williams & Wilkins, 1939.

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

Wechsler, D. Wechsler Adult Intelligence Scale. New York, The Psychological Corporation, 1955.

Wechsler, D. Wechsler Preschool and Primary Scale of Intelligence. New York, The Psychological Corporation, 1967.

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

REFERENCE NOTES

1. Catron, D. and Catron, S. WISC-R vs. WISC: A comparison with educable mentally retarded Children. Personal Communication, 1976.

2. Davis, C. Comparison of WISC and WISC-R performance scale for deaf students. in Swerdlik, M. The question of comparability of the WISC and WISC-R. Psychology in the Schools, July, 77. Vol. 14, No. 3.

3. Hartlage, Lawrence. The use of the WISC-R in the neuropsychological battery. Personal Communication, 1976.

4. Kaufman, A. and Weiner, S. A comparison of WISC-r and WISC for black children aged 7 to 10 years. Paper presented at the annual meeting of the Eastern Psychological Association, New York City, April 1976.

5. Solway, K., Fruge, E., Hays, J., Gryll, S., and Cody, J. Comparison of WISC and WISC-R scores for juvenile delinquents. Paper presented at the Annual Meeting of the Southwestern Psychological Association, Albuquerque, N.M., April 29-30, 1976.

6. Swerdlik, M. A comparison study of the WISC and WISC-R for children referred to school psychologists because of concerns about their intellectual ability. Unpublished doctoral dissertation, Michigan State University, August, 1976.

7. Swerdlik, M., and Rice, W. WISC minus WISC-R differences for referred children. Ingham Intermediate School District, Mason, Michigan, Summer, 1975.

8. Zimmerman, J. WISC-R minus WISC differences for referred children. Paper presented at the 1975 Annual Meeting of the American Psychological Association, Chicago, Illinois, September, 1975.