1998

Comparison of Ability-Achievement Discrepancies Among the Wechsler Individual Achievement Test, Woodcock-Johnson-Revised Tests of Achievement, and Wechsler Intelligence Scale for Children-Third Edition in a Sample of Students Referred for Evaluation of Learning Disabilities

Tracy L. Cole
Eastern Illinois University
This research is a product of the graduate program in School Psychology at Eastern Illinois University. Find out more about the program.

Recommended Citation
https://thekeep.eiu.edu/theses/1772

This is brought to you for free and open access by the Student Theses & Publications at The Keep. It has been accepted for inclusion in Masters Theses by an authorized administrator of The Keep. For more information, please contact tabruns@eiu.edu.
THESIS REPRODUCTION CERTIFICATE

TO: Graduate Degree Candidates (who have written formal theses)

SUBJECT: Permission to Reproduce Theses

The University Library is receiving a number of request from other institutions asking permission to reproduce dissertations for inclusion in their library holdings. Although no copyright laws are involved, we feel that professional courtesy demands that permission be obtained from the author before we allow these to be copied.

PLEASE SIGN ONE OF THE FOLLOWING STATEMENTS:

Booth Library of Eastern Illinois University has my permission to lend my thesis to a reputable college or university or the purpose of copying it for inclusion in that institution's library or research holdings.

7/24/98
Date

I respectfully request Booth Library of Eastern Illinois University NOT allow my thesis to be reproduced because:

_________________________________________________________________

_________________________________________________________________

_________________________________________________________________

_________________________________________________________________

_________________________________________________________________

Author's Signature ______________________________ Date ____________________
Comparison of Ability-Achievement Discrepancies Among the Wechsler

Individual Achievement Test, Woodcock-Johnson-Revised Tests of

Achievement, and Wechsler Intelligence Scale for Children-Third Edition in

a Sample of Students Referred for Evaluation of Learning Disabilities

BY

Tracy L. Cole

THESIS

SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF

Specialist in School Psychology

IN THE GRADUATE SCHOOL, EASTERN ILLINOIS UNIVERSITY
CHARLESTON, ILLINOIS

1998

YEAR

I HEREBY RECOMMEND THIS THESIS BE ACCEPTED AS FULFILLING
THIS PART OF THE GRADUATE DEGREE CITED ABOVE

23 July 1998
DATE

July 24, 1998
DATE
Abstract

Under current IDEA definitions, a significant or severe discrepancy between intellectual ability and academic achievement is the primary criterion by which a student may be classified as having a specific learning disability. By recommendation of the Work Group on Measurement Issues in the Assessment of Learning Disabilities (1984), the most common method of determining such a discrepancy is to calculate the discrepancy between an individual's score on an individually administered intelligence test and his/her score on an individually administered achievement test using a regression-based formula to correct for the regression of IQ on achievement.

Because two previous studies suggested that the Wechsler Individual Achievement Test (WIAT) produced lower scores in the areas of reading, mathematics, and writing than the Woodcock-Johnson-Revised Tests of Achievement (WJ-R ACH), the main focus of the current study was to determine whether the two achievement tests identify significant discrepancies consistently.

Data was collected anonymously from the files of 79 students who had been administered the WIAT, WJ-R ACH, and Wechsler Intelligence Scale for Children-Third Edition (WISC-III) as part of routine initial and reevaluations. Correlations, differences in mean standard scores, and discrepancy agreement statistics were calculated.
Correlations among WIAT and WJ-R ACH subtests and composites purporting to measure similar constructs yielded significant results. However, consistent with previous research, WIAT scores were significantly lower than WJ-R ACH scores in reading decoding, reading comprehension, reading composite, math reasoning, and written expression. The WJ-R ACH Dictation subtest produced significantly lower scores than the WIAT Spelling subtest. Furthermore, the WIAT produced significantly more discrepancies than the WJ-R ACH in the areas of Reading Comprehension, Written Expression, and Writing Composite.

The clinical implications, limitations and future directions are discussed.
Acknowledgements

I would like to thank my committee members, Dr. Gary Canivez, chairperson, Dr. Michael Havey, and Dr. Christine McCormick for their time and suggestions. My sincere thanks to Dr. Carl Hall, Anne Lewallen, Cynthia Hunt, and Sonja Castle for their professional concerns that necessitated this research and for their generosity in providing the data for this study. Finally, I would like to express my thanks to Brian and Samuel Cole for their patience, love, and assistance throughout this journey.
Discrepancies Between the WISC-III, WIAT, and WJ-R

Table of Contents

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Title Page</td>
<td>i</td>
</tr>
<tr>
<td>Abstract</td>
<td>1</td>
</tr>
<tr>
<td>Acknowledgements</td>
<td>3</td>
</tr>
<tr>
<td>Table of Contents</td>
<td>4</td>
</tr>
<tr>
<td>List of Tables</td>
<td>5</td>
</tr>
<tr>
<td>Introduction</td>
<td>6</td>
</tr>
<tr>
<td>Method</td>
<td>16</td>
</tr>
<tr>
<td>Participants</td>
<td>16</td>
</tr>
<tr>
<td>Instruments</td>
<td>16</td>
</tr>
<tr>
<td>Procedure</td>
<td>21</td>
</tr>
<tr>
<td>Data Analysis</td>
<td>21</td>
</tr>
<tr>
<td>Results</td>
<td>24</td>
</tr>
<tr>
<td>Discussion</td>
<td>35</td>
</tr>
<tr>
<td>References</td>
<td>49</td>
</tr>
<tr>
<td>Appendix</td>
<td>53</td>
</tr>
</tbody>
</table>
Discrepancies Between the WISC-III, WIAT, and WJ-R

List of Tables

<table>
<thead>
<tr>
<th>Table</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Table 1</td>
<td>Pearson Product-Moment Correlation Coefficients Between the WIAT and WJ-R ACH</td>
<td>25</td>
</tr>
<tr>
<td>Table 2</td>
<td>Pearson Product-Moment Correlation Coefficients Between the WISC-III and WIAT</td>
<td>27</td>
</tr>
<tr>
<td>Table 3</td>
<td>Pearson Product-Moment Correlation Coefficients Between the WISC-III and WJ-R ACH</td>
<td>28</td>
</tr>
<tr>
<td>Table 4</td>
<td>WIAT and WJ-R ACH Descriptive Statistics</td>
<td>30</td>
</tr>
<tr>
<td>Table 5</td>
<td>Discrepancy Agreement Statistics of Significant WISC-III-WIAT and WISC-III-WJ-R ACH Discrepancies Using WIAT Discrepancy Norms</td>
<td>33</td>
</tr>
<tr>
<td>Table 6</td>
<td>Discrepancy Agreement Statistics of Significant WISC-III-WIAT and WISC-III-WJ-R ACH Discrepancies Using WIAT-WISC-III Exact Correlations</td>
<td>34</td>
</tr>
<tr>
<td>Table 7</td>
<td>Discrepancy Agreement Statistics of Significant WISC-III-WIAT and WISC-III-WJ-R ACH Discrepancies Using Default Correlation of .65</td>
<td>35</td>
</tr>
</tbody>
</table>

According to the Individuals with Disabilities Education Act (IDEA; United States Department of Education [USDE], 1992), an individual may be classified as having a learning disability in any of the following seven areas: oral expression, listening comprehension, written expression, basic reading skill, reading comprehension, mathematics calculation, and mathematics reasoning. One criterion for the diagnosis of a learning disability is a significant discrepancy between intellectual ability and achievement (USDE, 1992, p. 44823). Since the passage of P.L. 94-142, the evaluation and diagnosis of learning disabilities has stirred much debate over several problems associated with the classification of the disability. Many of the problems stem from ambiguities in the definition of learning disabilities, as presented in P.L. 94-142. In this definition, classification of a learning disability:
is made based on (1) whether a child does not achieve commensurate with his or her age and ability when provided with appropriate educational experiences, and (2) whether the child has a severe discrepancy between achievement and intellectual ability in one or
more of seven areas relating to communication skills and mathematical abilities.

These concepts are to be interpreted on a case by case basis by the qualified evaluation team members. The team must decide that the discrepancy is not primarily the result of (1) visual, hearing, or motor handicaps; (2) mental retardation; (3) emotional disturbance; or (4) environmental, cultural, or economic disadvantage. (Federal Register, 1977, 42, p. 65082)

According to Reynolds (1984), a severe discrepancy between aptitude and achievement represents the only characteristic of a learning disability that is agreed upon in the field. However, the federal government did not provide criteria for determining a severe discrepancy in the final rules and regulations. Individual states were allowed to adopt and implement, within the parameters of the federal definition, criteria for diagnosing learning disabilities. Consequently, the models adopted from state to state for determining the existence of a severe discrepancy are varied and, at times, inadequate. In addition, the number of students diagnosed with a learning disability tripled from 1976 to 1982. It was in this context that USDE-SEP staff recommended that the Work Group on Measurement Issues in the Assessment of Learning Disabilities be formed to recommend best practices in determining severe discrepancies (Reynolds, 1984).
In 1984, Reynolds discussed the efforts of the Work Group to address the problems discussed above. The first issue addressed by the work group concerned the different measurement models and how they affect the proportion of children identified as learning disabled. Reynolds (1984) stated that, "the range of incidence figures easily can vary from less than 2% to more than 35% of a random sample of the population depending upon which state’s criteria are being applied" (p. 454).

Next, the Work Group discussed what types of children, under the different measurement models, are receiving LD services. The grade level discrepancy model, in which students performing two years below grade level are classified as LD, overidentifies students with IQ’s below 100 and underidentifies students with IQ’s above 100. Conversely, standard score comparison models that do not account for the regression of IQ on achievement will identify fewer children with IQ’s below 100 and more children with IQ’s over 100. This model mistakenly assumes that students should have an achievement level equal to that of their IQ. For example, a student with an IQ of 85 would be expected to have an achievement standard score of 85. However, due to regression, the achievement level of a student with an IQ of 85 would be about 88 or 89, depending on the correlation between the ability and achievement tests. Another model, the grade level exclusionary model, denies services to children who do not score below grade
level, regardless of IQ/achievement discrepancy. Application of this model will disqualify many students with IQ's over 100 for LD services.

In light of these models, the Work Group went on to describe what constitutes a severe discrepancy between aptitude and achievement from a statistical perspective. According to Reynolds (1984), "In determining what constitutes a severe discrepancy, the group consensus was that a regression model of some type must be adopted, that the simple difference score be reliable, and that the difference be relatively infrequent in the normal population" (p. 458). It was proposed that a discrepancy be considered infrequent at the .05 confidence level.

Necessary characteristics of the data used in determining a learning disability were also described. For instance, tests should be individually administered, unbiased, reliable, and standardized on a large, representative sample of the population. Standardization samples of aptitude and achievement measures on which children are being compared should consist of the same children (co-normed). All comparisons should be made using age-based standard scores, scaled to a common metric. Finally, any measures with subjective scoring criteria (e.g. written expression) should have interrater reliability of .85 to .90 or higher.

Scruggs & Mastropieri (1994), addressed many of the same issues as Reynolds (1984). In addition, the authors expressed concern that, "different tests of the same construct (e.g. achievement) can provide substantially
different results” (p. 20). Given the variety of individually administered tests for assessing academic achievement, whether different tests produce different results, with respect to severe discrepancies, is an important research question that will be explored in this study.

Stinnett, Havey, and Oehler-Stinnett (1994) randomly selected members of the National Association of School Psychologists (NASP) to complete surveys on which they rated the importance and frequency with which they used various testing instruments. Results showed that the Wechsler Intelligence Scale for Children-Third Edition (WISC-III; Wechsler, 1991) is one of the most frequently used instruments for the assessment of intellectual ability. In the area of academic achievement, the Woodcock-Johnson Psycho-Educational Battery-Revised Tests of Achievement (WJ-R ACH; Woodcock & Johnson, 1989) was the most frequently used instrument and received the highest mean importance rating in this area. Stinnett et al. (1994) asserted that the WJ-R ACH has sound psychometric properties. One reason for its frequent use could be that, “school psychologists probably are selecting psychometrically sound, standardized instruments and are less likely to use instruments with poor or unknown psychometric properties to assess academic achievement” (Stinnett et al., 1994, p. 343).

In 1992, the Wechsler Individual Achievement Test (WIAT) was published and normed on a subset of the standardization sample used in the standardization of the WISC-III. According to Flanagan & Alfonso (1993), “A
benefit of this link is the increased reliability that results from using co-
normed data to calculate discrepancies in the assessment and diagnosis of
learning disabilities” (p. 125). In addition, the WIAT is the first achievement
test designed to assess all seven areas of achievement (basic reading,
mathematics reasoning, reading comprehension, numerical operations,
listening comprehension, oral expression, written expression) in which a
learning disability can be diagnosed.

Given the popularity of the WJ-R ACH and the increasing popularity
of the WIAT among school psychologists, questions exist about differences
between the two instruments in the evaluation of learning disabilities. Do
these two achievement measures produce similar results?

As part of a series of criterion-related validity studies in the WIAT
Manual (1992), the WIAT and WJ-R ACH were administered to 43 children
aged 7 to 14 years. The correlations among subtests measuring the same
construct in the areas of reading, mathematics, and writing ranged from .74 -
.79, .67 - .68, and .72 - .88, respectively. Among the reading and mathematics
subtests, mean standard scores on the WIAT were 3.0 to 4.4 points lower than
WJ-R ACH standard scores. Conversely, mean WIAT standard scores on the
Written Expression and Spelling subtests were 9.2 and 1.6 points higher,
respectively, than scores on the Dictation subtest of the WJ-R ACH. The
authors noted that the WJ-R ACH Dictation subtest correlated higher with the
WIAT Spelling subtest ($r = .88$) than the WIAT Written Expression subtest
Discrepancies Between the WISC-III, WIAT, and WJ-R 12

$(r = .72)$. This pattern was expected due to the fact that the nature of the tasks on the Dictation and Spelling subtests are more similar than those of Written Expression.

Two limitations were inherent in this study. First, the Writing Samples subtest of the WJ-R ACH was not administered. It is possible that this particular subtest would have been more highly correlated with the WIAT’s Written Expression than was the Dictation subtest, given that the tasks required in the former more closely resemble those required in the Written Expression subtest. Another limitation of the study was that mean standard scores, standard deviations, and correlations were not reported for the composites. Because composite scores tend to be more reliable than individual subtest scores, this information would be valuable.

Martelle & Smith (1994) examined the relationship between the WIAT and the WJ-R ACH in a sample of 48 students referred for learning disability evaluations. Students were administered the WIAT and WJ-R ACH in counterbalanced order. The length of time between test administration ranged from one to 89 days, with an average of 15 days.

First, comparisons were made between WIAT and WJ-R ACH clusters (reading, mathematics, and writing) that appeared to measure the same construct. In the area of reading, significant correlations were found between the Reading Composite (RComp) of the WIAT and the Broad Reading (BR;
Discrepancies Between the WISC-III, WIAT, and WJ-R

$r = .70, \ p < .001$, Basic Reading Skills (BRS; $r = .81, \ p < .001$), and Reading Comprehension (RC; $r = .81, \ p < .005$) clusters of the WJ-R ACH. In all three comparisons, the WIAT scores were significantly ($p < .01$) lower than the WJ-R ACH scores by three to six points. In the area of mathematics, the Mathematics Composite (MComp) of the WIAT correlated significantly with the Broad Mathematics (BM; $r = .54, \ p < .001$) and Mathematics Reasoning (MR; $r = .43, \ p < .01$) clusters, but not with the Basic Mathematics Skills (BMS; $r = .14, \ NS$) cluster of the WJ-R ACH. Scores on the MComp were significantly lower than scores on the MR ($p < .001$) and BM ($p < .04$) clusters by three to seven points. The Language Composite of the WIAT was not significantly correlated with the Basic Written Language cluster of the WJ-R ACH ($r = .39$). Finally, in the area of writing, the Writing Composite (WComp) of the WIAT was moderately correlated with the Broad Written Language (BWL; $r = .59, \ p < .06$) cluster and was not significantly correlated with the Basic Writing Skills (BWS; $r = .12$) cluster of the WJ-R ACH.

WIAT/W mean scores were significantly lower ($p < .01$) than WJ-R/BWL mean scores by three points.

Martelle & Smith (1994) also examined relationships among individual subtests of the WIAT and WJ-R ACH. Reading Comprehension (RC) of the WIAT and Passage Comprehension (PC) of the WJ-R ACH were not significantly correlated. In the WIAT-RC subtest, students may use picture clues and/or key words in the passage to answer comprehension
questions. In the Passage Comprehension subtest, students read a passage and identify a missing word. Martelle & Smith explained that the two subtests differ in the approach to measuring reading comprehension and that these differences could account for the nonsignificant correlation.

Although most of the remaining subtests which appeared to be measuring the same construct were at least moderately correlated with one another, Martelle & Smith (1994) discussed differences among these subtests, as well. In the area of spelling, the WIAT subtest includes 11 homonyms, requiring the student to use their knowledge of the word in order to spell it correctly. The WJ-R ACH includes only two homonyms in the Spelling subtest. In the area of written language, the WIAT is less structured than that of the WJ-R ACH, allowing for more creativity and increased scorer judgment.

Martelle & Smith (1994) concluded that the WIAT and WJ-R ACH take different approaches to measuring academic skills. Consequently, the WIAT produces significantly lower scores in reading, mathematics, and written language than the WJ-R ACH by three to six points.

A limitation of this study was that mean scores of two subtests of the WJ-R ACH were not reported. While the administration of Letter-Word Identification and Dictation was implied through the derivation of Broad Reading and Broad Written Language cluster scores, the subtests were not
Discrepancies Between the WISC-III, WIAT, and WJ-R

referred to in text or tables. In addition, it would have been useful to analyze differences in mean scores for subtests measuring similar constructs.

As mentioned above, one criterion for the identification of a Specific Learning Disability is a severe discrepancy between intellectual ability, as estimated by a standardized intelligence test, and academic achievement, as estimated by a standardized achievement test. If, in fact, the findings of The Psychological Corporation (1992) and Martelle & Smith (1994) are considered from a statistical perspective, then it is possible that a larger number of significant discrepancies will be found between the WIAT and a given intelligence test than between the WJ-R ACH and the same intelligence test in the areas of reading, mathematics, and written language.

The focus of this study will be to examine the relationships among the WIAT, WJ-R ACH, and WISC-III by investigating the concurrent validity, construct validity, and diagnostic consistency of the tests. Specific research questions include: What are the correlations between the WIAT and WJ-R ACH at the subtest and global levels? When comparing mean standard scores, do the WIAT and WJ-R ACH tests produce similar results at the subtest and global levels? What are the correlations between the WIAT subtest and composite scores and the Full Scale IQ (FSIQ), Verbal IQ (VIQ), Performance IQ (PIQ), Verbal Comprehension Index (VCI), Perceptual Organization Index (POI), Freedom from Distractibility Index (FDI), and Processing Speed Index (PSI) obtained on the WISC-III? What are the
Discrepancies Between the WISC-III, WIAT, and WJ-R

Correlations between the WJ-R ACH subtest and cluster scores and FSIQ, VIQ, PIQ, VCI, POI, FDI, and PSI? When using the predicted-achievement method to determine a significant discrepancy from the FSIQ, will the WIAT and WJ-R ACH identify significant discrepancies consistently?

Method

Participants

Participants in this study included 79 students (35.4% female, 64.6% male) in grades K through 10, referred for initial (38%, n = 30) or reevaluation (62%, n = 49) for special education services. The students, 98.7% of whom were Caucasian, attended schools in a small, midwest city, public school district. Students ranged in age from 6 years, 4 months to 16 years, 6 months, with a mean age of 11 years, 3 months (SD = 2.73). The mean IQ of participants in the sample was 92.42 (SD = 13.51). More than half of the sample (53.2%; n = 42) was diagnosed with a learning disability, 5.1% (n = 4) was diagnosed with a mild mental impairment, and 41.8% (n = 33) were not found eligible to receive special education services.

Instruments

Discrepancies Between the WISC-III, WIAT, and WJ-R

Individually administered test of intellectual ability for use with children aged 6 through 16 years, 11 months. Three composite IQ scores, Verbal (VIQ), Performance (PIQ), and Full Scale (FSIQ), and four factor index scores, Verbal Comprehension (VCI), Perceptual Organization (POI), Freedom from Distractibility (FDI), and Processing Speed (PSI) are obtained for the WISC-III. The composite scores are derived from the scaled scores of ten subtests, five of which measure an individual's verbal/language abilities, and five of which measure visuo/spatial and problem-solving abilities. The Full Scale IQ is an estimation of an individual's general intellectual ability, while the VIQ and PIQ are estimates of verbal and nonverbal abilities, respectively.

The WISC-III was standardized on a representative sample of 2,200 students, with demographic data, such as gender, SES, race, ethnicity, and geographic region closely matching that of the 1988 U.S. Census. Internal consistency was calculated using the split-half method, with IQ and index score reliability coefficients ranging from .85 - .96. Test-retest stability for IQ and index scores ranged from .82 - .94. Practice effects increased the Full Scale IQ by 7 to 8 points at retest, which appeared to be more influenced by the Performance scale than by the Verbal scale. Interscorer reliability coefficients were in the high .90's for all subtests of the WISC-III except for Similarities, Vocabulary, Comprehension, and Mazes, which require some scorer judgment. For these subtests, interscorer reliability coefficients were .94, .92, .90, and .92, respectively. Factor analyses provided support for the four factor
structure of the WISC-III (Konold, Kush, & Canivez, 1997; Keith, 1994; Keith & Witta, 1994; Roid, et al., 1993; Wechsler, 1991). In addition, criterion-related validity was shown through correlations of the WISC-III and other measures of cognitive ability (r = .65 - .96) and with school grades (r = .47). Overall, the reliability and validity of the WISC-III are very good. Long term (three year) stability of FSIQ scores was also excellent (Canivez & Watkins, in press).

Wechsler Individual Achievement Test. The Wechsler Individual Achievement Test (WIAT; The Psychological Corporation, 1991) is an individually administered battery of achievement tests for students in grades K through 12 that assesses academic achievement in reading, mathematics, language, and writing. The eight subtests include Basic Reading, Mathematics Reasoning, Spelling, Reading Comprehension, Numerical Operations, Listening Comprehension, Oral Expression, and Written Expression. In addition, the WIAT yields composite scores in the areas of Reading, Mathematics, Language, and Writing. For the purposes of this study, the Listening Comprehension and Oral Expression subtests and the Language Composite were excluded because the WJ-R ACH does not have subtests or clusters pertain to oral expression or listening comprehension.

The WIAT was standardized on a representative sample of 4,252 children in grades K-12, with demographic data, such as gender, SES, race, ethnicity, and geographic region closely matching that of the 1988 U.S. Census. Of this sample, 1,284 students were also administered Wechsler
Intelligence Scales (WPPSI-R, WISC-III, WAIS-R). Internal consistency was calculated using the split-half method, with composite reliability coefficients ranging .88 - .97. Test-retest reliability coefficients ranged .65 - .97. Interscorer reliability was calculated for subtests requiring scorer judgment. Reading Comprehension and Listening Comprehension had an average interscorer reliability coefficient of .98, Oral Expression of .93, and Written Expression of .89 and .79 for Prompts 1 and 2, respectively. Content validity, which "refers to whether the items on a test are representative of the domain that the test purports to measure" (Sattler, 1992, p.30), was established using expert judgments and empirical item analysis. Construct validity, or the extent to which a test measures the construct or trait that it purports to measure, was evident through group differences in mean raw scores that follow similar patterns of other achievement tests. In addition, intercorrelations of subtests (reading with reading, math with math) provided evidence of discriminant validity, or the extent to which subtests measuring constructs in the same domain correlate more highly with each other than with subtests in other domains. Criterion-related validity was established through consistent correlations with other individually administered achievement tests measuring the same constructs. In addition, modest correlations were reported between WIAT scores and school grades. Overall, the reliability and validity of the WIAT are good.
Woodcock-Johnson Psycho-Educational Battery-Revised Tests of Achievement. The Woodcock-Johnson Psycho-Educational Battery-Revised Tests of Achievement (WJ-R ACH; Woodcock & Mather, 1989) is an individually administered battery of achievement tests that assesses academic achievement in reading, mathematics, writing, and general knowledge. Standard subtests of the WJ-R ACH include Letter-Word Identification, Passage Comprehension, Calculation, Applied Problems, Dictation, Writing Samples, Science, Social Studies, and Humanities. In addition, the WJ-R ACH yields cluster scores in the areas of Broad Reading, Broad Mathematics, Broad Written Language, Broad Knowledge, and Skills. For the purposes of this study, the Science, Social Studies, and Humanities subtests and the Broad Knowledge and Skills clusters were excluded because the WIAT does not have subtests or composites which pertain to these subject areas.

The WJ-R ACH was standardized on a representative sample of 6,359 subjects aged 2 through 90+ years, with demographic data closely approximating that of the 1980 U.S. Census. Internal consistency was calculated using the split-half method, with reliability coefficients in the mid .90's for the Broad achievement clusters. Test-retest and interscorer reliabilities were not reported in the Manual. Content validity was established through item validity studies and expert opinion. Concurrent validity correlations between the WJ-R ACH and other achievement measures at ages 3, 9, and 17 were in the .50's - .60's for age 3 and in the .60's
and .70's for ages 9 and 17. The fact that tests within curricular areas were more highly correlated than across curricular areas and that expected mean score patterns for special groups were obtained provided evidence of construct validity. Overall, the technical properties of the WJ-R ACH are adequate.

Procedure

Participants were administered the WISC-III, WIAT, and WJ-R ACH as part of psychoeducational evaluations within the school district. Data were obtained anonymously from existing student files. The examiners, three certified school psychologists, reported that order of administration varied but not through intentional counterbalancing.

Data Analysis

In order to investigate the convergent validity of the WIAT with the WJ-R ACH, Pearson product-moment correlation coefficients were calculated between WIAT and WJ-R ACH subtest and composite standard scores. In addition, WIAT and WJ-R ACH differences in mean standard scores between subtests and composites purported to measure similar constructs were analyzed. Because this was a two-group, within-subjects design and the dependent variable (standard scores) is measured on an interval scale, the appropriate statistical test for this purpose was dependent t-tests for differences between means. Previous research comparing the two tests is
limited; therefore, a two-tailed hypothesis test was used. In order to determine the strength of the difference between mean scores, effect sizes ($\eta^2$) were calculated using the formula (Keiss, 1996):

$$\eta^2 = \frac{t^2}{(t^2 + df)}$$

Pearson product-moment correlation coefficients were calculated between the WISC-III FSIQ, VIQ, PIQ, VCI, POI, FDI, and PSI and the WIAT and WJ-R ACH subtest and composite scores in order to examine the construct validity of the WIAT and WJ-R ACH with the WISC-III.

Another area of this study that was investigated was an examination of agreement or consistency of significant achievement-ability discrepancies identified using the WISC-III FSIQ and WIAT subtest and composite scores and discrepancies identified using the WISC-III FSIQ and WJ-R ACH subtest and cluster scores. Significant discrepancy was determined based on the predicted-achievement method using a regression formula. This method is recommended because it accounts for regression to the mean and error in measurement (Cone & Wilson, 1981; Evans, 1990; Heath & Kush, 1991; Reynolds, 1984).

Predicted achievement = $r_{xy} (IQ - M_{IQ}) = M_{ACH'}$, where IQ = the obtained IQ score, $M_{IQ} = 100$ (average IQ score), and $M_{ACH} = 100$ (average achievement...
Severe discrepancy between predicted achievement and actual achievement was defined by the formula:

\[ D > \sqrt{15}z (1 - r_{xy}) \]

(Reynolds, 1984), where \( D = \) Predicted Achievement-Actual Achievement, \( z = 1.65 \) (\( z \) corresponds to \( z = .05 \) in a one-tailed significance test), and \( r_{xy} = \) correlation between the ability and achievement measure. Because the correlation between the WISC-III and the WJ-R ACH for the general population is unknown, .65, the median IQ-Achievement correlation recommended by Heath & Kush (1991), was used. However, because the WISC-III and WIAT were co-normed, discrepancy norms, specific correlations (mean correlations between the WISC-III and WIAT for each subtest and composite), and the default correlation (\( r = .65 \)) were used in the calculation of a severe discrepancy. These values were obtained from the WIAT Manual. WISC-III Full Scale IQ was used as the IQ score for estimating achievement. Discrepancies were calculated on a computer by inserting the formula into a spreadsheet and not with a commercial program.

Diagnostic efficiency tables (Canivez & Watkins, 1996) were used to compare the presence or absence of severe achievement-ability discrepancies between the WIAT and WISC-III predicted achievement and the WJ-R ACH and WISC-III predicted achievement. (Diagnostic efficiency statistics were
calculated as recommended by Kessel & Zimmerman (1993). These statistics measure the extent to which the two tests agree in resulting classifications. In this study, discrepancy agreement tables were used to examine the overall agreement, or correct classification rate, between the WJ-R ACH and WIAT, with regard to severe discrepancy from the WISC-III.

Kappa, a coefficient of agreement beyond chance, was calculated to find the overall level of agreement. The general formula for kappa is:

\[ K = \frac{P_o - P_c}{1 - P_c} \]

where \( P_o \) = observed agreement and \( P_c \) = chance agreement. Kappa ranges from +1.00 to -1.00. A positive kappa indicates that observed agreement exceeds chance agreement and a negative kappa indicates that observed agreement is less than chance agreement (Sattler, 1992). Z-tests were calculated in order to determine the significance of kappa coefficients (Fleiss, 1981).

Results

Convergent validity between the WIAT and WJ-R ACH was examined by calculating Pearson product-moment correlation coefficients among the subtest and composite standard scores of the two tests. Correlations among
Discrepancies Between the WISC-III, WIAT, and WJ-R

subtests and composites purporting to measure constructs in similar domains (e.g. reading with reading) were all significant. See Table 1 for a complete listing of correlations among the subtests and composites.

Table 1
Pearson Product-Moment Correlation Coefficients Between the WIAT and WJ-R ACH

<table>
<thead>
<tr>
<th>WIAT</th>
<th>LWID</th>
<th>PC</th>
<th>C</th>
<th>AP</th>
<th>D</th>
<th>WS</th>
<th>Br-R</th>
<th>BM</th>
<th>BWL</th>
</tr>
</thead>
<tbody>
<tr>
<td>BR</td>
<td>.81**</td>
<td>.70**</td>
<td>.44**</td>
<td>.37*</td>
<td>.62**</td>
<td>.47**</td>
<td>.84**</td>
<td>.45**</td>
<td>.67**</td>
</tr>
<tr>
<td></td>
<td>(46)</td>
<td>(48)</td>
<td>(35)</td>
<td>(35)</td>
<td>(74)</td>
<td>(74)</td>
<td>(47)</td>
<td>(36)</td>
<td>(76)</td>
</tr>
<tr>
<td>MR</td>
<td>.28</td>
<td>.43**</td>
<td>.54**</td>
<td>.70**</td>
<td>.53**</td>
<td>.36**</td>
<td>.40**</td>
<td>.69**</td>
<td>.45**</td>
</tr>
<tr>
<td></td>
<td>(43)</td>
<td>(45)</td>
<td>(32)</td>
<td>(32)</td>
<td>(71)</td>
<td>(71)</td>
<td>(44)</td>
<td>(33)</td>
<td>(74)</td>
</tr>
<tr>
<td>S</td>
<td>.73*</td>
<td>.69**</td>
<td>.53**</td>
<td>.34*</td>
<td>.69**</td>
<td>.45**</td>
<td>.78**</td>
<td>.54**</td>
<td>.65**</td>
</tr>
<tr>
<td></td>
<td>(37)</td>
<td>(30)</td>
<td>(28)</td>
<td>(28)</td>
<td>(65)</td>
<td>(65)</td>
<td>(38)</td>
<td>(29)</td>
<td>(67)</td>
</tr>
<tr>
<td>RC</td>
<td>.70*</td>
<td>.74**</td>
<td>.45**</td>
<td>.57**</td>
<td>.53**</td>
<td>.43**</td>
<td>.77**</td>
<td>.54**</td>
<td>.53**</td>
</tr>
<tr>
<td></td>
<td>(46)</td>
<td>(48)</td>
<td>(35)</td>
<td>(35)</td>
<td>(74)</td>
<td>(74)</td>
<td>(47)</td>
<td>(36)</td>
<td>(76)</td>
</tr>
<tr>
<td>NO</td>
<td>.35</td>
<td>.43**</td>
<td>.60**</td>
<td>.81**</td>
<td>.62**</td>
<td>.35**</td>
<td>.44**</td>
<td>.75**</td>
<td>.53**</td>
</tr>
<tr>
<td></td>
<td>(43)</td>
<td>(45)</td>
<td>(32)</td>
<td>(32)</td>
<td>(71)</td>
<td>(71)</td>
<td>(44)</td>
<td>(33)</td>
<td>(74)</td>
</tr>
<tr>
<td>WE</td>
<td>.36</td>
<td>.33</td>
<td>.25</td>
<td>.14</td>
<td>.51**</td>
<td>.43*</td>
<td>.43</td>
<td>.29</td>
<td>.51**</td>
</tr>
<tr>
<td></td>
<td>(16)</td>
<td>(18)</td>
<td>(13)</td>
<td>(13)</td>
<td>(34)</td>
<td>(34)</td>
<td>(17)</td>
<td>(14)</td>
<td>(36)</td>
</tr>
<tr>
<td>R</td>
<td>.76*</td>
<td>.72**</td>
<td>.43**</td>
<td>.45**</td>
<td>.63**</td>
<td>.48**</td>
<td>.83**</td>
<td>.48**</td>
<td>.66**</td>
</tr>
<tr>
<td></td>
<td>(47)</td>
<td>(49)</td>
<td>(36)</td>
<td>(36)</td>
<td>(74)</td>
<td>(75)</td>
<td>(48)</td>
<td>(37)</td>
<td>(77)</td>
</tr>
<tr>
<td>M</td>
<td>.33</td>
<td>.39**</td>
<td>.58**</td>
<td>.77**</td>
<td>.60**</td>
<td>.39**</td>
<td>.44**</td>
<td>.73**</td>
<td>.51**</td>
</tr>
<tr>
<td></td>
<td>(44)</td>
<td>(46)</td>
<td>(33)</td>
<td>(33)</td>
<td>(72)</td>
<td>(72)</td>
<td>(45)</td>
<td>(34)</td>
<td>(75)</td>
</tr>
<tr>
<td>W</td>
<td>.75*</td>
<td>.64**</td>
<td>.63*</td>
<td>.35</td>
<td>.74**</td>
<td>.52**</td>
<td>.81**</td>
<td>.58*</td>
<td>.70**</td>
</tr>
<tr>
<td></td>
<td>(16)</td>
<td>(18)</td>
<td>(13)</td>
<td>(13)</td>
<td>(34)</td>
<td>(34)</td>
<td>(17)</td>
<td>(14)</td>
<td>(36)</td>
</tr>
</tbody>
</table>

Note.- WIAT = Wechsler Individual Achievement Test; WJ-R ACH = Woodcock-Johnson-Revised Tests of Achievement; BR = Basic Reading; MR = Mathematics Reasoning; S = Spelling; RC = Reading Comprehension; NO = Numerical Operations; WE = Written Expression; R = Reading Composite; M = Mathematics Composite; W = Writing Composite; LWID = Letter-Word Identification; PC = Passage Comprehension; C = Calculation; AP = Applied Problems; D = Dictation; WS = Writing Samples; Br-R = Broad Reading; BM = Broad Mathematics; BWL = Broad Written Language.

N presented in parentheses.

* p < .05  ** p < .01
Correlations among subtests and composites within similar domains ranged from .43 to .84, with a median correlation of .70. Correlations among subtests and composites in the reading domain ranged .70 to .84, with a median correlation of .76. In the area of math, correlations ranged from .54 to .81, with a median correlation of .70 and correlations in the writing domain ranged .43 to .74, with a median correlation of .52.

Among these correlations, the most meaningful are those purporting to measure similar constructs, including: BR/LWID, RC/PC, MR/AP, NO/C, S/D, WE/WS, R/Broad-R, M/BM, and W/BWL. Correlations for these comparisons ranged from .43 to .83, with a median correlation of .70.

Correlation coefficients were also calculated among WIAT and WJ-R ACH subtests and composites measuring constructs from different domains (reading with mathematics) in order to examine the discriminant validity of the tests. Resulting correlations ranged from .14 to .81, with a median correlation of .45.

In order to examine the construct validity of the WIAT and WJ-R ACH with the WISC-III, Pearson product-moment correlation coefficients were calculated among WISC-III IQ and index scores and WIAT and WJ-R ACH subtest and composite scores. Specific correlations can be found in Tables 2 & 3.

FSIQ correlations with WIAT subtest and composite standard scores ranged from .30 to .78, with a median correlation of .49. The verbal scales,
VIQ and VCI, produced correlation coefficients ranging .35 to .72, with a median correlation of .57. Correlations of the performance scales, PIQ and POI, with the WIAT ranged from .03 to .65, with a median correlation of .31.

Table 2
Pearson Product-Moment Correlation Coefficients Between the WISC-III and WIAT

<table>
<thead>
<tr>
<th>WIAT</th>
<th>VIQ</th>
<th>PIQ</th>
<th>FSIQ</th>
<th>VCI</th>
<th>POI</th>
<th>FDI</th>
<th>PSI</th>
</tr>
</thead>
<tbody>
<tr>
<td>BR</td>
<td>.41**</td>
<td>.03</td>
<td>.30**</td>
<td>.45**</td>
<td>.08</td>
<td>.36**</td>
<td>.06</td>
</tr>
<tr>
<td></td>
<td>(77)</td>
<td>(77)</td>
<td>(77)</td>
<td>(73)</td>
<td>(73)</td>
<td>(67)</td>
<td>(46)</td>
</tr>
<tr>
<td>MR</td>
<td>.70**</td>
<td>.56*</td>
<td>.78**</td>
<td>.68**</td>
<td>.65**</td>
<td>.64**</td>
<td>.21</td>
</tr>
<tr>
<td></td>
<td>(75)</td>
<td>(75)</td>
<td>(75)</td>
<td>(71)</td>
<td>(71)</td>
<td>(65)</td>
<td>(46)</td>
</tr>
<tr>
<td>S</td>
<td>.36**</td>
<td>.16</td>
<td>.34**</td>
<td>.35**</td>
<td>.21</td>
<td>.37**</td>
<td>.31</td>
</tr>
<tr>
<td></td>
<td>(68)</td>
<td>(68)</td>
<td>(68)</td>
<td>(64)</td>
<td>(64)</td>
<td>(59)</td>
<td>(39)</td>
</tr>
<tr>
<td>RC</td>
<td>.65**</td>
<td>.27*</td>
<td>.58**</td>
<td>.68**</td>
<td>.36**</td>
<td>.56**</td>
<td>.05</td>
</tr>
<tr>
<td></td>
<td>(77)</td>
<td>(77)</td>
<td>(77)</td>
<td>(73)</td>
<td>(73)</td>
<td>(67)</td>
<td>(46)</td>
</tr>
<tr>
<td>NO</td>
<td>.67**</td>
<td>.51**</td>
<td>.71**</td>
<td>.62**</td>
<td>.54**</td>
<td>.63**</td>
<td>.41**</td>
</tr>
<tr>
<td></td>
<td>(75)</td>
<td>(75)</td>
<td>(75)</td>
<td>(71)</td>
<td>(71)</td>
<td>(65)</td>
<td>(46)</td>
</tr>
<tr>
<td>WE</td>
<td>.41*</td>
<td>.26</td>
<td>.39*</td>
<td>.38*</td>
<td>.25</td>
<td>.34</td>
<td>.07</td>
</tr>
<tr>
<td></td>
<td>(37)</td>
<td>(37)</td>
<td>(37)</td>
<td>(35)</td>
<td>(35)</td>
<td>(31)</td>
<td>(21)</td>
</tr>
<tr>
<td>R</td>
<td>.55**</td>
<td>.13</td>
<td>.44**</td>
<td>.59**</td>
<td>.20</td>
<td>.46**</td>
<td>.06</td>
</tr>
<tr>
<td></td>
<td>(78)</td>
<td>(78)</td>
<td>(78)</td>
<td>(74)</td>
<td>(74)</td>
<td>(68)</td>
<td>(47)</td>
</tr>
<tr>
<td>M</td>
<td>.72**</td>
<td>.56**</td>
<td>.78**</td>
<td>.68**</td>
<td>.63**</td>
<td>.67**</td>
<td>.32*</td>
</tr>
<tr>
<td></td>
<td>(76)</td>
<td>(76)</td>
<td>(76)</td>
<td>(72)</td>
<td>(72)</td>
<td>(66)</td>
<td>(47)</td>
</tr>
<tr>
<td>W</td>
<td>.50**</td>
<td>.35*</td>
<td>.50**</td>
<td>.41*</td>
<td>.34*</td>
<td>.44*</td>
<td>.46*</td>
</tr>
<tr>
<td></td>
<td>(37)</td>
<td>(37)</td>
<td>(37)</td>
<td>(35)</td>
<td>(35)</td>
<td>(31)</td>
<td>(21)</td>
</tr>
</tbody>
</table>

Note.- WISC-III = Wechsler Intelligence Scale for Children-Third Edition; VIQ = Verbal IQ; PIQ = Performance IQ; FSIQ = Full Scale IQ; VCI = Verbal Comprehension Index; POI = Perceptual Organization Index; FDI = Freedom from Distractibility Index; POI = Processing Speed Index; WIAT = Wechsler Individual Achievement Test; BR = Basic Reading; MR = Mathematics Reasoning; S = Spelling; RC = Reading Comprehension; NO = Numerical Operations; WE = Written Expression; R = Reading Composite; M = Mathematics Composite; W = Writing Composite.

N presented in parentheses.
* p < .05  ** p < .01
Discrepancies Between the WISC-III, WIAT, and WJ-R  28

FDI and PSI correlations ranged .05 to .67, with median correlations of .46 and .21, respectively.

Table 3
Pearson Product-Moment Correlation Coefficients Between the WISC-III and WJ-R ACH

<table>
<thead>
<tr>
<th>WJ-R ACH</th>
<th>VIQ</th>
<th>PIQ</th>
<th>FSIQ</th>
<th>VCI</th>
<th>POI</th>
<th>FDI</th>
<th>PSI</th>
</tr>
</thead>
<tbody>
<tr>
<td>LWID</td>
<td>.20 (.47)</td>
<td>-.05 (.47)</td>
<td>.16 (.47)</td>
<td>.20 (.44)</td>
<td>.09 (.44)</td>
<td>.37* (.41)</td>
<td>.15 (.30)</td>
</tr>
<tr>
<td>PC</td>
<td>.37** (.49)</td>
<td>.14 (.49)</td>
<td>.35* (.49)</td>
<td>.41** (.46)</td>
<td>.21 (.46)</td>
<td>.35* (.43)</td>
<td>.11 (.31)</td>
</tr>
<tr>
<td>C</td>
<td>.49** (.36)</td>
<td>.28 (.36)</td>
<td>.60** (.36)</td>
<td>.55** (.933)</td>
<td>.55** (.33)</td>
<td>.48** (.31)</td>
<td>.40 (.24)</td>
</tr>
<tr>
<td>AP</td>
<td>.63** (.36)</td>
<td>.47** (.36)</td>
<td>.71** (.936)</td>
<td>.58** (.933)</td>
<td>.58** (.33)</td>
<td>.60** (.31)</td>
<td>.52* (.24)</td>
</tr>
<tr>
<td>D</td>
<td>.45** (.75)</td>
<td>.18 (.75)</td>
<td>.44** (.75)</td>
<td>.45** (.72)</td>
<td>.31** (.72)</td>
<td>.48** (.66)</td>
<td>.18 (.47)</td>
</tr>
<tr>
<td>WS</td>
<td>.36** (.75)</td>
<td>.16 (.75)</td>
<td>.35** (.75)</td>
<td>.41** (.72)</td>
<td>.26* (.72)</td>
<td>.12 (.66)</td>
<td>-.04 (.47)</td>
</tr>
<tr>
<td>Br-R</td>
<td>.31* (.48)</td>
<td>.05 (.48)</td>
<td>.28 (.48)</td>
<td>.31* (.44)</td>
<td>.15 (.44)</td>
<td>.46** (.41)</td>
<td>.22 (.30)</td>
</tr>
<tr>
<td>BM</td>
<td>.60** (.37)</td>
<td>.40* (.37)</td>
<td>.70** (.37)</td>
<td>.60** (.33)</td>
<td>.60** (.33)</td>
<td>.59** (.31)</td>
<td>.55** (.24)</td>
</tr>
<tr>
<td>BWL</td>
<td>.43** (.78)</td>
<td>.18 (.78)</td>
<td>.42** (.78)</td>
<td>.47** (.74)</td>
<td>.50** (.74)</td>
<td>.30* (.68)</td>
<td>.06 (.48)</td>
</tr>
</tbody>
</table>

Note: WISC-III = Wechsler Intelligence Scale for Children-Third Edition; VIQ = Verbal IQ; PIQ = Performance IQ; FSIQ = Full Scale IQ; VCI = Verbal Comprehension Index; POI = Perceptual Organization Index; FDI = Freedom from Distractibility Index; POI = Processing Speed Index; WJ-R ACH = Woodcock-Johnson-Revised Tests of Achievement; LWID = Letter-Word Identification; PC = Passage Comprehension; C = Calculation; AP = Applied Problems; D = Dictation; WS = Writing Samples; Br-R = Broad Reading; BM = Broad Mathematics; BWL = Broad Written Language.

N presented in parentheses.

* p < .05  ** p < .01
Discrepancies Between the WISC-III, WIAT, and WJ-R

FSIQ correlations with WJ-R ACH subtest and cluster standard scores ranged from .16 to .71, with a median correlation of .42. The verbal scales yielded correlation coefficients ranging .20 to .63, with a median correlation of .44. Correlations of the performance scales with the WJ-R ACH ranged from -.05 to .60, with a median correlation of .24. FDI and PSI correlations ranged from -.04 to .60, with median correlations of .46 and .18, respectively.

Given the significant correlations between WIAT and WJ-R ACH subtest and composite scores purporting to measure similar constructs, differences in mean standard scores of WIAT and WJ-R ACH subtests and composites were analyzed to determine if the two tests produce similar results. Two-tailed dependent t-tests were calculated between the six subtests and three composites of the WIAT with corresponding subtests and clusters of the WJ-R ACH. Results are presented in Table 4.

In the area of reading, the WIAT produced significantly lower scores than the WJ-R ACH at the subtest and composite levels. WIAT/BR (M = 86.85) was significantly lower than WJ-R ACH/LWID (M = 88.96); WIAT/RC (M = 88.44) was significantly lower than WJ-R ACH/PC (M = 92.94); and WIAT/R (M = 84.92) was significantly lower than WJ-R ACH/Broad-R (M = 90.38). While the effect sizes for RC/PC and R/Broad-R were moderate, the effect size for BR/LWID was low.
Table 4
WIAT and WJ-R ACH Descriptive Statistics

<table>
<thead>
<tr>
<th>Composites</th>
<th>WIAT M</th>
<th>SD</th>
<th>Clusters</th>
<th>WJ-R M</th>
<th>SD</th>
<th>t</th>
<th>df</th>
<th>p</th>
<th>n²</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reading</td>
<td>84.92</td>
<td>10.29</td>
<td>Broad-R</td>
<td>90.38</td>
<td>11.30</td>
<td>-5.87</td>
<td>47</td>
<td>.000</td>
<td>.42</td>
</tr>
<tr>
<td>Math</td>
<td>85.06</td>
<td>10.99</td>
<td>BM</td>
<td>88.59</td>
<td>15.15</td>
<td>-2.00</td>
<td>33</td>
<td>.054</td>
<td>.11</td>
</tr>
<tr>
<td>Writing</td>
<td>86.06</td>
<td>12.52</td>
<td>BWL</td>
<td>88.58</td>
<td>9.39</td>
<td>0.32</td>
<td>35</td>
<td>.752</td>
<td>.00</td>
</tr>
<tr>
<td>Subtests</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BR</td>
<td>86.85</td>
<td>10.10</td>
<td></td>
<td>88.96</td>
<td>11.74</td>
<td>-2.06</td>
<td>45</td>
<td>.045</td>
<td>.09</td>
</tr>
<tr>
<td>MR</td>
<td>88.63</td>
<td>10.40</td>
<td>AP</td>
<td>94.63</td>
<td>13.72</td>
<td>-3.45</td>
<td>31</td>
<td>.002</td>
<td>.28</td>
</tr>
<tr>
<td>S</td>
<td>85.83</td>
<td>10.72</td>
<td>D</td>
<td>82.80</td>
<td>8.78</td>
<td>3.12</td>
<td>64</td>
<td>.003</td>
<td>.13</td>
</tr>
<tr>
<td>RC</td>
<td>88.44</td>
<td>8.82</td>
<td>PC</td>
<td>92.94</td>
<td>11.23</td>
<td>-4.13</td>
<td>47</td>
<td>.000</td>
<td>.27</td>
</tr>
<tr>
<td>NO</td>
<td>84.38</td>
<td>9.72</td>
<td>C</td>
<td>84.72</td>
<td>16.26</td>
<td>-0.15</td>
<td>31</td>
<td>.882</td>
<td>.00</td>
</tr>
<tr>
<td>WE</td>
<td>84.74</td>
<td>17.30</td>
<td>WS</td>
<td>91.47</td>
<td>13.11</td>
<td>-2.37</td>
<td>33</td>
<td>.024</td>
<td>.15</td>
</tr>
</tbody>
</table>

Note.- WIAT = Wechsler Individual Achievement Test; BR = Basic Reading; MR = Mathematics Reasoning; S = Spelling; RC = Reading Comprehension; NO = Numerical Operations; WE = Written Expression; WJ-R ACH = Woodcock-Johnson-Revised Tests of Achievement; Broad-R = Broad Reading; BM = Broad Mathematics; BWL = Broad Written Language; LWID = Letter-Word Identification; AP = Applied Problems; D = Dictation; PC = Passage Comprehension; C = Calculation; WS = Writing Samples.

In the area of math, the WIAT and WJ-R ACH produced more consistent results than in the reading domain. WIAT/MR (M = 88.63) was significantly lower than WJ-R ACH/AP (M = 94.63), with a moderate effect size. However, WIAT/NO (M = 84.38) did not yield results significantly different from that of the WJ-R ACH/C (M = 84.72); Overall, the difference between the math composite scores approached significance, with the
Discrepancies Between the WISC-III, WIAT, and WJ-R

WIAT/M ($M = 85.06$) producing somewhat lower scores than the WJ-R ACH/BM ($M = 88.59$).

In the writing domain, WIAT and WJ-R ACH composite scores were similar, while subtest scores were significantly different. The WIAT/S ($M = 85.83$) was significantly higher than WJ-R ACH/D ($M = 82.80$), while WIAT/WE ($M = 84.74$) was significantly lower than WJ-R ACH/WS ($M = 91.47$). Overall, the WIAT/W ($M = 86.06$) was not significantly different from WJ-R ACH/BWL ($M = 85.58$). Effect sizes for all three comparisons were low.

Because significant differences were found among mean standard scores for six of the nine comparisons, overall agreement or consistency of significant achievement-ability discrepancies identified using the WISC-III FSIQ and WIAT versus WJ-R ACH subtest and composite scores was examined. Significant discrepancies were determined based on the predicted-achievement method using the regression method (Reynolds, 1984). Significant discrepancies between the WISC-III and WIAT were calculated using the following three methods: 1) WISC-III-WIAT Discrepancy Norms based on the co-normed standardization sample (The Psychological Corporation, 1992); 2) calculation of a regression formula using WISC-III-WIAT specific correlations based on the co-normed standardization sample (The Psychological Corporation, 1992); and 3) calculation of a regression formula using a default correlation of .65, the median IQ-Achievement correlation recommended by Heath & Kush (1991, p. 9). Because specific
Discrepancies Between the WISC-III, WIAT, and WJ-R

correlations between the WISC-III and WJ-R ACH are yet unknown, the
default correlation of .65 was also used to calculate the regression based
discrepancy between the WISC-III and WJ-R ACH.

Using each method of calculating a significant discrepancy between the
WISC-III and WIAT, discrepancy agreement tables (Canivez & Watkins, 1996)
were used to compare the presence or absence of significant discrepancies
between the WIAT and WISC-III predicted achievement with the presence or
absence of significant discrepancies between the WJ-R ACH and WISC-III
predicted achievement. Discrepancy Agreement tables for each comparison
can be found in the Appendix. These tables present the overall agreement
between the WIAT and WJ-R ACH, with regard to significant discrepancy
from the WISC-III. In order to find the overall level of agreement, kappa, a
coefficient of agreement beyond chance was calculated. Statistical significance
of kappa was also obtained. Tables 5, 6, and 7 present the discrepancy
agreement statistics for each of the three WISC-III-WIAT methods.

Analysis of discrepancy agreement comparing the WIAT discrepancy
norms to the WJ-R ACH default correlation produced agreement ranging
from .57 to .91. However, with chance agreement ranging from .58 to .91,
kappa was significant for only three comparisons: BR/LWID, S/D, and
R/Broad-R. Kappa coefficients ranged -.24 to .49.

An examination of discrepancy agreement between the WIAT specific
and default correlations and WJ-R ACH default correlation resulted in
agreement ranging from .57 to 1.00. Chance agreement ranged .60 to 1.00, with significant kappa found for BR/LWID, S/D, WE/WS, R/Broad-R, and W/BWL. Kappa could not be calculated for MR/AP due to limited variability in the sample.

Table 5

<table>
<thead>
<tr>
<th>Subtests</th>
<th>Percent Agreement</th>
<th>Chance Agreement</th>
<th>Kappa</th>
<th>SE&lt;sub&gt;k&lt;/sub&gt;</th>
<th>Z</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>BR/LWID</td>
<td>.83</td>
<td>.67</td>
<td>.49</td>
<td>.14</td>
<td>3.46</td>
<td>.001</td>
</tr>
<tr>
<td>MR/AP</td>
<td>.91</td>
<td>.91</td>
<td>0</td>
<td>.01</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>S/D</td>
<td>.78</td>
<td>.67</td>
<td>.33</td>
<td>.12</td>
<td>2.68</td>
<td>.007</td>
</tr>
<tr>
<td>RC/PC</td>
<td>.57</td>
<td>.58</td>
<td>-.04</td>
<td>.10</td>
<td>-.37</td>
<td>.708</td>
</tr>
<tr>
<td>NO/C</td>
<td>.58</td>
<td>.66</td>
<td>-.24</td>
<td>.16</td>
<td>-1.49</td>
<td>.137</td>
</tr>
<tr>
<td>WE/WS</td>
<td>.76</td>
<td>.70</td>
<td>.21</td>
<td>.16</td>
<td>1.28</td>
<td>.199</td>
</tr>
<tr>
<td>R/Broad-R</td>
<td>.73</td>
<td>.64</td>
<td>.25</td>
<td>.12</td>
<td>2.00</td>
<td>.045</td>
</tr>
<tr>
<td>M/BM</td>
<td>.74</td>
<td>.78</td>
<td>-.15</td>
<td>.17</td>
<td>-.87</td>
<td>.386</td>
</tr>
<tr>
<td>W/BWL</td>
<td>.78</td>
<td>.69</td>
<td>.29</td>
<td>.16</td>
<td>1.79</td>
<td>.073</td>
</tr>
</tbody>
</table>

Note.- WISC-III = Wechsler Intelligence Scale for Children-Third Edition; WIAT = Wechsler Individual Achievement Test; WJ-R ACH = Woodcock-Johnson-Revised Tests of Achievement; BR = Basic Reading; MR = Mathematics Reasoning; S = Spelling; RC = Reading Comprehension; NO = Numerical Operations; WE = Written Expression; R = Reading Composite; M = Mathematics Composite; W = Writing Composite; LWID = Letter-Word Identification; PC = Passage Comprehension; C = Calculation; AP = Applied Problems; D = Dictation; WS = Writing Samples; Broad-R = Broad Reading; BM = Broad Mathematics; BWL = Broad Written Language.
Discrepancies Between the WISC-III, WIAT, and WJ-R

Table 6

<table>
<thead>
<tr>
<th>Subtests</th>
<th>Percent Agreement</th>
<th>Chance Agreement</th>
<th>Kappa</th>
<th>$SE_k$</th>
<th>Z</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>BR/LWID</td>
<td>.85</td>
<td>.73</td>
<td>.44</td>
<td>.15</td>
<td>3.02</td>
<td>.003</td>
</tr>
<tr>
<td>MR/AP</td>
<td>1.00</td>
<td>1.00</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>S/D</td>
<td>.85</td>
<td>.72</td>
<td>.45</td>
<td>.12</td>
<td>3.65</td>
<td>.001</td>
</tr>
<tr>
<td>RC/PC</td>
<td>.57</td>
<td>.60</td>
<td>-.08</td>
<td>.08</td>
<td>-1.13</td>
<td>.259</td>
</tr>
<tr>
<td>NO/C</td>
<td>.81</td>
<td>.83</td>
<td>-.09</td>
<td>.16</td>
<td>-0.55</td>
<td>.581</td>
</tr>
<tr>
<td>WE/WS</td>
<td>.82</td>
<td>.74</td>
<td>.32</td>
<td>.15</td>
<td>2.07</td>
<td>.019</td>
</tr>
<tr>
<td>R/Broad-R</td>
<td>.79</td>
<td>.69</td>
<td>.33</td>
<td>.13</td>
<td>2.52</td>
<td>.012</td>
</tr>
<tr>
<td>M/BM</td>
<td>.79</td>
<td>.81</td>
<td>-.11</td>
<td>.17</td>
<td>-0.66</td>
<td>.507</td>
</tr>
<tr>
<td>W/BWL</td>
<td>.86</td>
<td>.78</td>
<td>.37</td>
<td>.17</td>
<td>2.21</td>
<td>.027</td>
</tr>
</tbody>
</table>

Note.- WISC-III = Wechsler Intelligence Scale for Children-Third Edition; WIAT = Wechsler Individual Achievement Test; WJ-R ACH = Woodcock-Johnson-Revised Tests of Achievement; BR = Basic Reading; MR = Mathematics Reasoning; S = Spelling; RC = Reading Comprehension; NO = Numerical Operations; WE = Written Expression; R = Reading Composite; M = Mathematics Composite; W = Writing Composite; LWID = Letter-Word Identification; PC = Passage Comprehension; C = Calculation; AP = Applied Problems; D = Dictation; WS = Writing Samples; Broad-R = Broad Reading; BM = Broad Mathematics; BWL = Broad Written Language. Missing entries designate inability to calculate due to limited variability in scores.
Table 7
Discrepancy Agreement Statistics of Significant WISC-III-WIAT and WISC-III-WJ-R ACH Discrepancies Using Default Correlation of .65

<table>
<thead>
<tr>
<th>Subtests</th>
<th>Percent Agreement</th>
<th>Chance Agreement</th>
<th>Kappa</th>
<th>SE\textsubscript{\kappa}</th>
<th>Z</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>BR/LWID</td>
<td>.85</td>
<td>.70</td>
<td>.50</td>
<td>.15</td>
<td>3.37</td>
<td>.001</td>
</tr>
<tr>
<td>MR/AP</td>
<td>1.00</td>
<td>1.00</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>S/D</td>
<td>.73</td>
<td>.64</td>
<td>.26</td>
<td>.11</td>
<td>2.41</td>
<td>.016</td>
</tr>
<tr>
<td>RC/PC</td>
<td>.57</td>
<td>.60</td>
<td>-.08</td>
<td>.08</td>
<td>-1.13</td>
<td>.259</td>
</tr>
<tr>
<td>NO/C</td>
<td>.81</td>
<td>.83</td>
<td>-.09</td>
<td>.16</td>
<td>-0.55</td>
<td>.581</td>
</tr>
<tr>
<td>WE/WS</td>
<td>.82</td>
<td>.74</td>
<td>.32</td>
<td>.15</td>
<td>2.07</td>
<td>.039</td>
</tr>
<tr>
<td>R/Broad-R</td>
<td>.79</td>
<td>.69</td>
<td>.33</td>
<td>.13</td>
<td>2.52</td>
<td>.012</td>
</tr>
<tr>
<td>M/BM</td>
<td>.79</td>
<td>.81</td>
<td>-.11</td>
<td>.17</td>
<td>-0.66</td>
<td>.507</td>
</tr>
<tr>
<td>W/BWL</td>
<td>.86</td>
<td>.78</td>
<td>.37</td>
<td>.17</td>
<td>2.21</td>
<td>.027</td>
</tr>
</tbody>
</table>

Note. - WISC-III = Wechsler Intelligence Scale for Children-Third Edition; WIAT = Wechsler Individual Achievement Test; WJ-R ACH = Woodcock-Johnson-Revised Tests of Achievement; BR = Basic Reading; MR = Mathematics Reasoning; S = Spelling; RC = Reading Comprehension; NO = Numerical Operations; WE = Written Expression; R = Reading Composite; M = Mathematics Composite; W = Writing Composite; LWID = Letter-Word Identification; PC = Passage Comprehension; C = Calculation; AP = Applied Problems; D = Dictation; WS = Writing Samples; Broad-R = Broad Reading; BM = Broad Mathematics; BWL = Broad Written Language. Missing entries designate inability to calculate due to limited variability in scores.

Discussion

The passage of IDEA (USDE, 1992) created the opportunity for individuals with a specific learning disability, as defined in the law, to receive special education services. The primary component of the definition, the existence of a significant discrepancy between intellectual ability and academic
Discrepancies Between the WISC-III, WIAT, and WJ-R achievement, has been interpreted differently from state to state and the models for determining such a discrepancy have been varied.

In 1984, by the recommendation of USDE-SEP staff, the Work Group on Measurement Issues in the Assessment of Learning Disabilities discussed the various diagnostic models and recommended best practices in determining what constitutes a significant discrepancy. Due to statistical inadequacies, the grade level discrepancy model, the standard score comparison model, and the grade level exclusionary model were deemed inappropriate by the Work Group. The consensus was that a model that accounts for the regression of IQ on achievement be used to determine a severe discrepancy.

In order to determine a significant discrepancy, individually administered achievement and IQ measures are given. In 1994, a study by Stinnett, et. al., reported that the WJ-R ACH was the most frequently used achievement test. However, the 1992 publication of the WIAT, which was co-normed with the popular WISC-III, provided the increased reliability and convenience of using co-normed data to determine discrepancies.

An article by Scruggs & Mastropieri (1994) argued that different tests, in this case the WJ-R ACH and WIAT, produce different results. Previous studies (Psychological Corporation, 1992; Martelle & Smith, 1994) suggested that the WIAT and WJ-R ACH yield different scores. In the areas of reading, mathematics, and written language, the WIAT produced lower scores than the WJ-R ACH.
Discrepancies Between the WISC-III, WIAT, and WJ-R

Given the regression model for determining a significant discrepancy, as recommended by the Work Group (1984), the WIAT would be likely, statistically, to produce more significant discrepancies than the WJ-R ACH. This poses a problem for special educators and school psychologists who use these tests for making important educational decisions. The purpose of this study was to examine the relationships among the WIAT, WJ-R ACH, and WISC-III and to investigate the extent to which the two achievement tests produce similar results with respect to significant discrepancy from the WISC-III.

The first investigation involved the convergent validity, or the extent to which two tests purporting to measure similar constructs, namely the WIAT and WJ-R ACH, are correlated. Pearson product-moment correlations were calculated among all subtest and composite standard scores of the WIAT and WJ-R ACH. In the areas of reading, mathematics, and writing, all correlations were significant at the subtest and composite levels. Correlations in the three areas ranged .70 - .84, .54 - .81, and .43 - .74, respectively. As expected, correlations between composites in similar domains were among the highest, ranging from .70 to .83.

Divergent validity was also examined using Pearson product-moment correlations. Within the three domains, most of the subtests purporting to measure similar constructs (e.g. word identification with word identification) were more highly correlated with each other than with subtests measuring a
different construct (e.g. word identification with reading comprehension).
WIAT/BR was more highly correlated with WJ-R ACH/LWID than with WJ-R ACH/PC; WIAT/RC was more highly correlated with WJ-R ACH/PC than with WJ-R ACH/LWID; WIAT/MR was more highly correlated with WJ-R ACH/AP than with WJ-R ACH/C; and WIAT/S was more highly correlated with WJ-R ACH/D than with WJ-R ACH/WS. However, WIAT/NO did not correlate as highly with WJ-R ACH/C as it did with WJ-R ACH/AP and WIAT/WE did not correlate as highly with WJ-R ACH/WS as it did with WJ-R ACH/D.

Pearson product-moment correlations among WIAT and WJ-R ACH subtest and composite scores yielded significant results, providing evidence for the convergent validity of the two achievement tests for the current study sample. Given the significant correlations between tests of similar constructs, differences in mean standard scores of WIAT and WJ-R ACH subtests and composites were examined using t-tests for dependent means.

Consistent with previous studies comparing the WIAT and WJ-R ACH (Martelle & Smith, 1994; The Psychological Corporation, 1992), significant differences in mean scores were found in six of the nine comparisons. The WIAT yielded significantly lower scores than the WJ-R ACH in reading decoding, reading comprehension, and the reading composite by two to five points. In addition, WIAT Math Reasoning and Written Expression were significantly lower than the WJ-R ACH Applied Problems and Writing
Samples by six and seven points, respectively. Conversely, the WJ-R ACH Dictation subtest was significantly lower, by an average of three points, than the WIAT Spelling subtest. WIAT Numerical Operations, Writing Composite, and Math Composite yielded similar results to the corresponding subtest and clusters of the WJ-R ACH.

Differences in mean scores were similar between the present study and previous studies. At the subtest level, results were similar between the current study and the validity study conducted by the Psychological Corporation (1992), which found WIAT/NO to be significantly lower than WJ-R ACH/C and WIAT/S not to differ significantly from WJ-R ACH/D. At the global level, the only difference between Martelle and Smith (1994) and the present study was that Martelle and Smith (1994) found significantly lower scores on the WIAT/W than on the WJ-R ACH/BWL.

In the area of reading, while word recognition (WIAT Basic Reading and WJ-R ACH Letter-Word Identification) is measured in a similar way, reading comprehension is assessed differently. The Reading Comprehension subtest of the WIAT requires the students to answer a comprehension question pertaining to a written passage. On the WJ-R ACH Passage Comprehension subtest, a student reads a short passage and produces a word that will make sense of a sentence with an omission. Students must give a short phrase or sentence to receive credit on the WIAT/RC, while one word is sufficient for
credit on the WJ-R ACH/PC. Length of response may contribute to the
disparity in mean standard scores in the area of reading comprehension.

In the area of math, the WIAT and WJ-R ACH appear to measure math
reasoning and math calculation similarly. This is evidenced by the similar
standard scores produced by the two achievement tests in math calculations
and overall math composite.

Both components of the writing domain, spelling and written
expression, produced different scores on the WIAT and WJ-R ACH. Not
surprisingly, the approach to measuring the two constructs differs between the
two tests. For instance, as noted in Martelle & Smith (1994), the WIAT Spelling
subtest contains eleven homonyms, while the WJ-R ACH Dictation subtest
includes only two homonyms. Therefore, many of the items on the WIAT
require the student to use their knowledge of the word in order to spell it
correctly. In addition, while the WIAT/S focuses only on spelling, the WJ-R
ACH/D contains several items pertaining to word usage, punctuation, and
capitalization. The Spelling and Dictation subtests clearly tap different skills
associated with the technical aspects of writing, resulting in incongruent scores.
Many school psychologists believe that a significant discrepancy in Spelling or
Dictation, alone, is not sufficient for the diagnosis of a specific learning
disability.

The other component within the writing domain, written expression, is
also assessed differently between the WIAT and WJ-R ACH. The Written
Expression subtest of the WIAT requires the student to write a letter to a friend based on one of two prompts chosen by the examiner. The composition is then rated on a scale of one to four by the examiner in the following six areas: ideas and development; organization, unity and coherence; vocabulary; sentence structure and variety; grammar and usage; capitalization and punctuation. The WJ-R ACH Writing Samples subtest requires the individual to write a phrase or sentence for each item. Each test item is rated from zero to two by the examiner based on the quality of the response. Errors in spelling and punctuation are not penalized. The criteria for scoring the Writing Samples are more clearly defined than those for scoring Written Expression. While the Written Expression subtest more closely approximates the writing demands of a classroom setting than the Writing Samples subtest, the scoring is more subjective and open to scorer judgment differences.

Five subtests and one composite of the WIAT produced scores that were significantly different (higher or lower) than those of the WJ-R ACH. These differences could affect discrepancy calculations, as lower scores are more likely to result in significant discrepancies from intellectual ability scores. Because significant differences were found between six of the nine relevant comparisons of the WIAT and WJ-R ACH, it would follow that agreement in the number of significant discrepancies from the WISC-III FSIQ should be examined. However, before investigating this area, the relationship between
the two achievement tests and the ability measure with which they are being compared was examined.

Pearson product-moment correlation coefficients were calculated among WIAT and WJ-R ACH subtest and composite scores and WISC-III IQ and index scores in order to determine if the relationships between the WIAT and WISC-III are similar to the relationships between the WJ-R ACH and WISC-III. Overall, the Verbal IQ and Verbal Comprehension Index produced the highest correlations with achievement scores of both tests. The Processing Speed Index score had the lowest correlations with achievement scores. Furthermore, verbal measures, VIQ and VCI, correlated with achievement scores as well as or better than did FSIQ. Conversely, nonverbal scales, PIQ and POI, yielded lower correlations with achievement scores than did VIQ, VCI, or FSIQ. According to Kaufman & Kaufman (1990) and Wechsler (1991), these relationships would be expected.

Using the regression method to calculate predicted achievement, significant discrepancies between the WJ-R ACH subtest and cluster scores and the WISC-III FSIQ were determined in only one way, using a default correlation of .65, because exact correlations with the WISC-III for the general population are still unknown. However, predicted achievement for the WIAT was calculated in each of three possible ways. The most popular, convenient, and technically appropriate (Reynolds, 1984) method of calculating an individual’s predicted achievement is to use the discrepancy norms tables
located in the Appendix of the WIAT Manual. Another method is to use a regression formula to calculate predicted achievement using exact correlations. Finally, a regression formula was calculated using the default correlation of .65.

For each comparison, kappa, a coefficient of agreement beyond chance, was calculated to find the overall level of agreement. In order for kappa to yield a valid coefficient, there must be at least one case in each of the four cells of the discrepancy agreement table (see Appendix). Due to a limited sample size, some comparisons did not have enough variability among scores to produce a valid measure of kappa, even though agreement may have been 100%. For example, on one comparison, both the WIAT and WJ-R ACH determined that none of the 32 cases were significantly discrepant from the WISC-III FSIQ. Although this constituted perfect agreement, kappa could not be validly calculated with the remaining three cells of the table left empty.

In the area of reading, the WIAT discrepancy norms, specific correlations, and default correlations produced results similar to those obtained with the WJ-R ACH. Kappa, the coefficient of agreement beyond chance, was significant for WIAT/BR and WJ-R ACH/LWID and for WIAT/R and WJ-R ACH/Broad-R. This means that there was significant agreement between the WIAT and WJ-R ACH in the areas of word recognition and overall reading. However, kappa was not significant for WIAT/RC and WJ-R ACH/PC. In other words, the WIAT produced
significantly more discrepancies than the WJ-R ACH in the area of reading comprehension.

In the area of math, observed agreement was high, ranging from .58 to 1.00. However, the study sample consisted of a limited number of students suspected of having a learning disability in the math areas. Therefore, with limited variability, chance agreement equaled or exceeded observed agreement, resulting in nonsignificant kappa coefficients for all comparisons in this area.

In the writing domain, kappa was significant for all comparisons using the WIAT/WISC-III exact correlations and the default correlation. However, when using the WIAT discrepancy norms, the WIAT/WE and W identified significantly more discrepancies than the WJ-R ACH/WS and BWL.

In some comparisons, while the WIAT produced significantly lower scores than the WJ-R ACH, significant kappas indicated agreement with regard to severe discrepancies. However, an examination of the Discrepancy Agreement tables in the Appendix suggest that when there is disagreement between the two tests, the WIAT produces more discrepancies than the WJ-R ACH. In other words, the WJ-R ACH yields a higher rate of false negatives.

When comparing WIAT discrepancies across all three methods, a larger number of discrepancies were identified when using the WIAT discrepancy norms. McGrew, Werder, and Woodcock (1991) discuss the difference between discrepancy norms and regression based methods for
calculating significant discrepancies. When calculating discrepancy norms, predicted-achievement is calculated using actual scores from the co-normed achievement and ability measures. When achievement and intellectual measures are not co-normed, the next best method for determining a significant discrepancy is to employ a regression based method. In this method, predicted achievement is "estimated, not calculated, using a correction for regression to the mean" (McGrew, Werder, & Woodcock, 1991, p.84). Therefore, in this study, the discrepancy norms method identified significant discrepancies based on actual predicted-achievement, whereas the other two methods determined discrepancy based on estimated predicted-achievement.

Results of the analysis of discrepancy agreement indicates that, in general, the WIAT and WJ-R ACH identify significant discrepancies consistently in all areas except reading comprehension when using the WIAT/WISC-III exact correlations and the default correlation. The WIAT also produced significantly more discrepancies in Written Expression and the Writing Composite than did the WJ-R ACH when comparing the WIAT discrepancy norms with the WJ-R ACH default correlation. Therefore, significant differences in mean standard scores affected the identification of significant discrepancies in two of the six comparisons. In addition, the WIAT discrepancy norms yielded significantly more discrepancies on the Writing Composite than on the WJ-R ACH/BWL even though mean
standard scores were not found to be significantly different. If the full battery of each achievement test were given in an evaluation, a student would be more likely to evidence a significant discrepancy between intellectual ability and academic achievement when comparing the WIAT to the WISC-III predicted-achievement than when comparing the WJ-R ACH to the WISC-III predicted-achievement.

The clinical implications of this study are significant for special educators and school psychologists. Results indicated that in psychoeducational evaluations (initial and reevaluation), choosing the WIAT over the WJ-R ACH, or vice versa, could affect whether a student is identified as having a significant discrepancy between intellectual ability and academic achievement. This, in turn, affects whether a student is diagnosed with a specific learning disability and subsequent qualification for special educational programming. Because the WIAT and WISC-III were co-normed, as suggested by the Work Group, using the WIAT in the assessment of learning disabilities would be optimal.

As with any standardized test, error in measurement is inherent in the WIAT, WJ-R ACH, and WISC-III, which can produce false positives and false negatives with regard to significant discrepancies. An example of a false positive would be a student who exhibits a significant discrepancy between achievement and ability, whose achievement is truly not discrepant. This might cause a multidisciplinary team to qualify this student for special
education services when she truly does not qualify. Conversely, a false negative would be a case where a student whose achievement is truly discrepant from his ability does not demonstrate a significant discrepancy. This case could cause a multidisciplinary team to deny special education services to a student who truly qualifies for such services.

The major limitations of this study are associated with the sample. First, while the total number of participants in the study was adequate (n = 79), not every participant was administered the full battery of both achievement tests. Because the students were administered the tests as part of a routine psychoeducational evaluation, many participants received only the tests related to his or her suspected area of disability. Therefore, sample sizes for individual comparisons ranged from 32 to 65, with a median sample size of 36 within comparisons. Data analyses, particularly the calculation of kappa, was limited by the small size and homogeneity of the sample. As discussed previously, all participants were suspected of having a learning disability, limiting the variability among cells of the discrepancy agreement tables. In addition, of the 79 participants, 78 were Caucasian. Therefore, the sample for this study was not representative of the larger national or state populations and findings should not be generalized to minority groups.

Given the clinical implications, as well as the limitations, of this study, further research comparing the WIAT and WJ-R ACH would be useful. Inclusion of minority students and students demonstrating average to above
average academic performance, as well as students with other disabilities, would facilitate the reliable and valid calculation of kappa when examining the overall level of agreement between the WIAT and WJ-R ACH with regard to significant discrepancy from the WISC-III.

Another area of needed research relating to the current study is an examination of the relationship between the WJ-R ACH and WISC-III. Administration of the WJ-R ACH and WISC-III to a large, representative sample would allow better estimates of the correlations between these two popular tests to be obtained for the general population. This information would have been useful for inclusion in this study because a more valid comparison could have been made between WIAT/WJ-R ACH discrepancies from the WISC-III. In addition, such correlations would be beneficial to school psychologists in reliably and validly computing a regression based discrepancy formula to identify a significant discrepancy between intellectual ability and academic achievement.
Discrepancies Between the WISC-III, WIAT, and WJ-R

References


Discrepancies Between the WISC-III, WIAT, and WJ-R


Discrepancies Between the WISC-III, WIAT, and WJ-R  53

Appendix
Discrepancy Agreement

Table

WIAT/BR (discrepancy norms) by
W-R ACH/LWID

<table>
<thead>
<tr>
<th>WJ-R ACH</th>
<th>Not Discrepant</th>
<th>Discrepant</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>WIAT</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Not Discrepant</td>
<td>32</td>
<td>2</td>
<td>34</td>
</tr>
<tr>
<td>Discrepant</td>
<td>6</td>
<td>6</td>
<td>12</td>
</tr>
<tr>
<td>Total</td>
<td>38</td>
<td>8</td>
<td>46</td>
</tr>
</tbody>
</table>

Results

Overall Agreement (Hit) Rate = .8261

Observed Agreement Po = .8261
Chance Agreement Pc = .656

Kappa = .4945
Standard Error of Kappa = .142741187

Significance Test for Kappa
Ho: k = 3.464311947
0  Z =

p < .00053168  two-tail test
p < .0026584  one-tail test

© 1994 by Gary L. Canivez, Ph.D.
All rights reserved.
### Discrepancy Agreement Table

WIAT/BR (specific correlations)  
by W-R ACH/LWID

<table>
<thead>
<tr>
<th>WJ-R Ach</th>
<th>Not Discrepant</th>
<th>Discrepant</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>WIAT</td>
<td>35</td>
<td>4</td>
<td>39</td>
</tr>
<tr>
<td>Discrepant</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>4</td>
<td>7</td>
<td></td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>38</strong></td>
<td><strong>8</strong></td>
<td><strong>46</strong></td>
</tr>
</tbody>
</table>

### Results

Overall Agreement (Hit) Rate = .8478  
Observed Agreement Po = .8478  
Chance Agreement Pc = .7268

Kappa = .4429  
Standard Error of Kappa = .146846259

Significance Test for Kappa  
Ho: k = 3.016079552  
\[ Z = \]  
\[ p < .0025608 \]  
\[ p < .0012804 \]  
two-tail test  
one-tail test

© 1994 by Gary L. Canivez, Ph.D.  
All rights reserved.
**Discrepancy Agreement Table**

WIAT/BR (default correlations) by W-R ACH/LWID

<table>
<thead>
<tr>
<th></th>
<th>WIAT Not Discrepant</th>
<th>Discrepant</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Not Discrepant</td>
<td>34</td>
<td>3</td>
<td>37</td>
</tr>
<tr>
<td>Discrepant</td>
<td>4</td>
<td>5</td>
<td>9</td>
</tr>
<tr>
<td>Total</td>
<td>38</td>
<td>8</td>
<td>46</td>
</tr>
</tbody>
</table>

**Results**

- Overall Agreement (Hit) Rate = .8478
- Observed Agreement Po = .8478
- Chance Agreement Pc = .6985
- Kappa = .4952
- Standard Error of Kappa = .14708816
- Significance Test for Kappa Ho: k = 3.366688386
- \( Z = 0 \)
- \( p < .00076087 \) two-tail test
- \( p < .00038044 \) one-tail test

© 1994 by Gary L. Canivez, Ph.D.
All rights reserved.
Discrepancies Between the WISC-III, WIAT, and WJ-R

**Discrepancy Agreement Table**

WIAT/RC (discrepancy norms) by W-R ACH/PC

<table>
<thead>
<tr>
<th>WIAT Not Discrepant</th>
<th>Discrepant</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>39</td>
<td>25</td>
<td>64</td>
</tr>
<tr>
<td>8</td>
<td>4</td>
<td>12</td>
</tr>
<tr>
<td>Total</td>
<td>47</td>
<td>29</td>
</tr>
</tbody>
</table>

**Results**

Overall Agreement (Hit) Rate = .5658

Observed Agreement Po = .5658
Chance Agreement Pc = .581

Kappa = -.0363
Standard Error of Kappa = .096965502

Significance Test for Kappa Ho: k = -.374359944 0  Z =

p < .70813665 two-tail test
p < .545406832 one-tail test

© 1994 by Gary L. Canivez, Ph.D.
All rights reserved.
**Discrepancy Agreement**

*Table*

WIAT/RC (specific correlations)  
by W-R ACH/PC

<table>
<thead>
<tr>
<th>WJ-R Ach</th>
<th>Not Discrepant</th>
<th>Discrepant</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>WIAT</td>
<td>42</td>
<td>28</td>
<td>70</td>
</tr>
<tr>
<td>Not Discrepant</td>
<td>5</td>
<td>1</td>
<td>6</td>
</tr>
<tr>
<td>Discrepant</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>47</td>
<td>29</td>
<td>76</td>
</tr>
</tbody>
</table>

**Results**

Overall Agreement (Hit) Rate = .5658

Observed Agreement Po = .5658  
Chance Agreement Pc = .5997

Kappa = -.0847  
Standard Error of Kappa = .075045206

Significance Test for Kappa  
Ho: k = -1.128653045  
0  
Z =  
p < .25904432  
p < .12952216  
two-tail test  
one-tail test

© 1994 by Gary L. Canivez, Ph.D.  
All rights reserved.
Discrepancy Agreement

Table

WIAT/RC (default correlations) by W-R ACH/PC

<table>
<thead>
<tr>
<th>WIAT/RC</th>
<th>Not Discrepant</th>
<th>Discrepant</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Not Discrepant</td>
<td>42</td>
<td>28</td>
<td>70</td>
</tr>
<tr>
<td>Discrepant</td>
<td>5</td>
<td>1</td>
<td>6</td>
</tr>
<tr>
<td>Total</td>
<td>47</td>
<td>29</td>
<td>76</td>
</tr>
</tbody>
</table>

Results

Overall Agreement (Hit) Rate = .5658

Observed Agreement Po = .5658
Chance Agreement Pc = .5997

Kappa = -.0847
Standard Error of Kappa = .075045206

Significance Test for Kappa

Ho: k = -1.128653045

Z =

p < .25904432 two-tail test
p < .12952216 one-tail test
Discrepancies Between the WISC-III, WIAT, and WJ-R

**Discrepancy Agreement**

**Table**

WIAT/R (discrepancy norms) by W-R ACH/Broad-R

<table>
<thead>
<tr>
<th>WIAT Not Discrepant</th>
<th>Discrepant</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>31</td>
<td>2</td>
<td>33</td>
</tr>
<tr>
<td>11</td>
<td>4</td>
<td>15</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>42</strong></td>
<td><strong>48</strong></td>
</tr>
</tbody>
</table>

**Results**

Overall Agreement (Hit) Rate = .7292

Observed Agreement Po = .7292

Chance Agreement Pc = .6406

Kappa = .2465

Standard Error of Kappa = .123089188

Significance Test for Kappa

Ho: k = 2.002612938

Z =

p < .04521871 two-tail test

p < .02260936 one-tail test

© 1994 by Gary L. Canivez, Ph.D.
All rights reserved.
**Discrepancy Agreement Table**

WIAT/R (specific correlations) by W-R ACH/Broad-R

<table>
<thead>
<tr>
<th>WIAT Not Discrepant</th>
<th>Discrepant</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>36</td>
</tr>
<tr>
<td>WIAT Not Discrepant</td>
<td></td>
<td></td>
</tr>
<tr>
<td>34</td>
<td>2</td>
<td>36</td>
</tr>
<tr>
<td>WIAT Discrepant</td>
<td></td>
<td>12</td>
</tr>
<tr>
<td>8</td>
<td>4</td>
<td>12</td>
</tr>
<tr>
<td>Total</td>
<td>42</td>
<td>6</td>
</tr>
</tbody>
</table>

**Results**

Overall Agreement (Hit) Rate = .7917

Observed Agreement Po = .7917
Chance Agreement Pc = .6875

Kappa = .3334
Standard Error of Kappa = .132287566

Significance Test for Kappa
Ho: k = 2.520267106

0  Z =

p < .01172662   two-tail test
p < .00586331   one-tail test
Discrepancies Between the WISC-III, WIAT, and WJ-R  62

**Discrepancy Agreement Table**

WIAT/R (default correlations) by W-R ACH/Broad-R

<table>
<thead>
<tr>
<th>WJ-R ACH</th>
<th>Not Discrepant</th>
<th>Discrepant</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>WIAT</td>
<td>34</td>
<td>2</td>
<td>36</td>
</tr>
<tr>
<td>Not Discrepant</td>
<td>8</td>
<td>4</td>
<td>12</td>
</tr>
</tbody>
</table>

**Results**

Overall Agreement (Hit) Rate = .7917

Observed Agreement Po = .7917
Chance Agreement Pc = .6875

Kappa = .3334
Standard Error of Kappa = .132287566

Significance Test for Kappa Ho: k = 2.520267106
0  Z =
p < .01172662 two-tail test
p < .00586331 one-tail test

© 1994 by Gary L. Canivez, Ph.D.
All rights reserved.
**Discrepancy Agreement Table**

WIAT/MR (discrepancy norms) by W-R ACH/AP

<table>
<thead>
<tr>
<th>WJ-R Ach</th>
<th>Not Discrepant</th>
<th>Discrepant</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>WIAT</td>
<td>30</td>
<td>0</td>
<td>30</td>
</tr>
<tr>
<td>Discrepant</td>
<td>3</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>Total</td>
<td>33</td>
<td>0</td>
<td>33</td>
</tr>
</tbody>
</table>

**Results**

- Overall Agreement (Hit) Rate = .9091
- Observed Agreement Po = .9091
- Chance Agreement Pc = .9091
- Kappa = 0
- Standard Error of Kappa = .009693219

Significance Test for Kappa

\[ H_0: k = 0 \]

\[ Z = \]

\[ p < 1 \quad \text{two-tail test} \]

\[ p < .5 \quad \text{one-tail test} \]
**Discrepancy Agreement Table**

WIAT/MR (specific correlations)  
by W-R ACH/AP

<table>
<thead>
<tr>
<th></th>
<th>WIAT Not Discrepant</th>
<th>WJ-R Ach Discrepant</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>WIAT Not Discrepant</td>
<td>32</td>
<td>0</td>
<td>32</td>
</tr>
<tr>
<td>WIAT Discrepant</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>32</td>
<td>0</td>
<td>32</td>
</tr>
</tbody>
</table>

**Results**

<table>
<thead>
<tr>
<th>Description</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overall Agreement (Hit) Rate</td>
<td>1</td>
</tr>
<tr>
<td>Observed Agreement Po</td>
<td>1</td>
</tr>
<tr>
<td>Chance Agreement Pc</td>
<td>1</td>
</tr>
<tr>
<td>Kappa = #DIV/O!</td>
<td></td>
</tr>
<tr>
<td>Standard Error of Kappa = #DIV/O!</td>
<td></td>
</tr>
<tr>
<td>Significance Test for Kappa Ho: k</td>
<td>#DIV/O!</td>
</tr>
<tr>
<td>0 Z =</td>
<td></td>
</tr>
<tr>
<td>p &lt; #DIV/O! two-tail test</td>
<td></td>
</tr>
<tr>
<td>p &lt; #DIV/O! one-tail test</td>
<td></td>
</tr>
</tbody>
</table>

© 1994 by Gary L. Canivez, Ph.D.  
All rights reserved.
### WJ-R Ach

<table>
<thead>
<tr>
<th>WIAT</th>
<th>Not Discrepant</th>
<th>Discrepant</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Not</td>
<td>32</td>
<td>0</td>
<td>32</td>
</tr>
<tr>
<td>Discrepant</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Total</td>
<td>32</td>
<td>0</td>
<td>32</td>
</tr>
</tbody>
</table>

**Results**

Overall Agreement (Hit) Rate = 1

- Observed Agreement $P_o = 1$
- Chance Agreement $P_c = 1$

$$Kappa = \frac{#DIV/O!}{#DIV/O!}$$

- Standard Error of Kappa $= \frac{#DIV/O!}{#DIV/O!}$

Significance Test for Kappa
- $H_0: k = #DIV/O!$
- $Z = \frac{0}{#DIV/O!}$
- $p < #DIV/O!$ two-tail test
- $p < #DIV/O!$ one-tail test

© 1994 by Gary L. Canivez, Ph.D.
All rights reserved.
Discrepancy Agreement Table

WIAT/NO (discrepancy norms) by W-R ACH/C

<table>
<thead>
<tr>
<th>WJ-R Ach</th>
<th>Not Discrepant</th>
<th>Discrepant</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>WIAT</td>
<td>19</td>
<td>5</td>
<td>24</td>
</tr>
<tr>
<td>Not</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Discrepant</td>
<td>9</td>
<td>0</td>
<td>9</td>
</tr>
<tr>
<td>Total</td>
<td>28</td>
<td>5</td>
<td>33</td>
</tr>
</tbody>
</table>

Results

Overall Agreement (Hit) Rate = .5758

Observed Agreement Po = .5758
Chance Agreement Pc = .6584

Kappa = -.2418
Standard Error of Kappa = .16274487

Significance Test for Kappa
Ho: k = -1.485761121
0 Z =

p < .13734238 two-tail test
p < .06867119 one-tail test

© 1994 by Gary L. Canivez, Ph.D.
All rights reserved.
Discrepancies Between the WISC-III, WIAT, and WJ-R

Discrepancy Agreement Table

WIAT/NO (specific correlations) by W-R ACH/C

<table>
<thead>
<tr>
<th>WJ-R Ach</th>
<th>Not Discrepant</th>
<th>Discrepant</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>WIAT</td>
<td>26</td>
<td>4</td>
<td>30</td>
</tr>
<tr>
<td>Not Discrepant</td>
<td>2</td>
<td>0</td>
<td>2</td>
</tr>
</tbody>
</table>

Total 28 4 32

Results

Overall Agreement (Hit) Rate = .8125

Observed Agreement Po = .8125
Chance Agreement Pc = .8281

Kappa = -.0908
Standard Error of Kappa = .16443788

Significance Test for Kappa Ho: k = -.552185856
0 Z =
p < .58082096 two-tail test
p < .29041048 one-tail test

© 1994 by Gary L. Canivez, Ph.D.
All rights reserved.
**Discrepancy Agreement Table**

WIAT/NO (default correlations) by W-R ACH/C

<table>
<thead>
<tr>
<th>WIAT Not Discrepant</th>
<th>Discrepant</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>26</td>
<td>4</td>
<td>30</td>
</tr>
<tr>
<td>2</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>28</strong></td>
<td><strong>32</strong></td>
</tr>
</tbody>
</table>

**Results**

Overall Agreement (Hit) Rate = .8125

Observed Agreement \( P_o \) = .8125  
Chance Agreement \( P_c \) = .8281

Kappa = -.0908  
Standard Error of Kappa = .164437388

Significance Test for Kappa  
Ho: \( k = -.552185856 \)

\[ Z = \]
\[ p < .58082096 \text{ (two-tail test)} \]
\[ p < .29041048 \text{ (one-tail test)} \]
Discrepancy Agreement Table

WIAT/M (discrepancy norms) by W-R ACH/BM

<table>
<thead>
<tr>
<th>WJ-R Ach</th>
<th>Not Discrepant</th>
<th>Discrepant</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>WIAT Not Discrepant</td>
<td>26</td>
<td>4</td>
<td>30</td>
</tr>
<tr>
<td>WIAT Discrepant</td>
<td>5</td>
<td>0</td>
<td>5</td>
</tr>
<tr>
<td>Total</td>
<td>31</td>
<td>4</td>
<td>35</td>
</tr>
</tbody>
</table>

Results

Overall Agreement (Hit) Rate = 0.7429

Observed Agreement Po = 0.7429

Chance Agreement Pc = 0.7755

Kappa = -0.1452

Standard Error of Kappa = 0.167604615

Significance Test for Kappa Ho: k = -0.866324594

0  Z =

p < 0.38631209 two-tail test

p < 0.19315605 one-tail test

© 1994 by Gary L. Canivez, Ph.D.
All rights reserved.
Discrepancy Agreement

Table

WIAT/M (specific correlations) by W-R ACH/BM

<table>
<thead>
<tr>
<th>WIAT Not Discrepant</th>
<th>WJ-R ACH</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Not Discrepant</td>
<td></td>
<td>27</td>
</tr>
<tr>
<td>Discrepant</td>
<td></td>
<td>3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>30</td>
</tr>
<tr>
<td>Discrepant</td>
<td></td>
<td>4</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>4</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>31</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>34</td>
</tr>
</tbody>
</table>

Results

Overall Agreement (Hit) Rate = .7941

Observed Agreement Po = .7941
Chance Agreement Pc = .8149

Kappa = -.1124
Standard Error of Kappa = .169480176

Significance Test for Kappa Ho: k = -.663204407
0 Z =

p < .50719951 two-tail test
p < .25359975 one-tail test

© 1994 by Gary L. Canivez, Ph.D.
All rights reserved.
Discrepancies Between the WISC-III, WIAT, and WJ-R

**Discrepancy Agreement**

**Table**

WIAT/M (default correlations) by W-R ACH/BM

<table>
<thead>
<tr>
<th>WJ-R Ach</th>
<th>Not Discrepant</th>
<th>Discrepant</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>WIAT</td>
<td>27</td>
<td>3</td>
<td>30</td>
</tr>
<tr>
<td>Not Discrepant</td>
<td>4</td>
<td>0</td>
<td>4</td>
</tr>
</tbody>
</table>

Total 31 3 34

**Results**

Overall Agreement (Hit) Rate = .7941

Observed Agreement Po = .7941
Chance Agreement Pc = .8149

Kappa = -.1124
Standard Error of Kappa = .169480176

Significance Test for Kappa

\[ Z = \frac{-0.663204407}{0.169480176} \]

\[ p < .00719951 \] two-tail test
\[ p < .25359975 \] one-tail test

© 1994 by Gary L. Canivez, Ph.D.
All rights reserved.
Discrepancies Between the WISC-III, WIAT, and WJ-R

**Discrepancy Agreement Table**

WIAT/S (discrepancy norms) by W-R ACH/D

<table>
<thead>
<tr>
<th>WJ-R Ach</th>
<th>Not Discrepant</th>
<th>Discrepant</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>WIAT</td>
<td>44</td>
<td>5</td>
<td>49</td>
</tr>
<tr>
<td>Not Discrepant</td>
<td>9</td>
<td>6</td>
<td>15</td>
</tr>
<tr>
<td>Discrepant</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>53</td>
<td>11</td>
<td>64</td>
</tr>
</tbody>
</table>

**Results**

Overall Agreement (Hit) Rate = .7813

- Observed Agreement Po = .7813
- Chance Agreement Pc = .6743

- Kappa = .3285
- Standard Error of Kappa = .122647397

Significance Test for Kappa

\[ Z = \frac{2.678409887 - 0}{0.122647397} \]

- \( p < .00739736 \) two-tail test
- \( p < .00369868 \) one-tail test

© 1994 by Gary L. Canivez, Ph.D.
All rights reserved.
Discrepancy Agreement

Table

WIAT/S (specific correlations) by W-R ACH/D

<table>
<thead>
<tr>
<th></th>
<th>Not Discrepant</th>
<th>Discrepant</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>WIAT Not Discrepant</td>
<td>49</td>
<td>5</td>
<td>54</td>
</tr>
<tr>
<td>WIAT Discrepant</td>
<td>5</td>
<td>6</td>
<td>11</td>
</tr>
<tr>
<td>Total</td>
<td>54</td>
<td>11</td>
<td>65</td>
</tr>
</tbody>
</table>

Results

Overall Agreement (Hit) Rate = .8462

Observed Agreement Po = .8462
Chance Agreement Pc = .7188

Kappa = .4531
Standard Error of Kappa = .123995745

Significance Test for Kappa Ho: k = 3.654157635
0 Z = p < .00025809 two-tail test
     p < .00012905 one-tail test

© 1994 by Gary L. Canivez, Ph.D.
All rights reserved.
Discrepancies Between the WISC-III, WIAT, and WJ-R

**Discrepancy Agreement**

**Table**

WIAT/S (default correlations) by W-R ACH/D

<table>
<thead>
<tr>
<th></th>
<th>WIAT Not Discrepant</th>
<th>Discrepant</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>WJ-R Ach</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Not Discrepant</td>
<td>48</td>
<td>5</td>
<td>53</td>
</tr>
<tr>
<td>Discrepant</td>
<td>15</td>
<td>7</td>
<td>22</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>63</strong></td>
<td><strong>12</strong></td>
<td><strong>75</strong></td>
</tr>
</tbody>
</table>

**Results**

Overall Agreement (Hit) Rate = .7333

- Observed Agreement Po = .7333
- Chance Agreement Pc = .6405

Kappa = .2581

Standard Error of Kappa = .107186454

Significance Test for Kappa

- Ho: k = 2.407953517
- 0 Z =
  - p < .01604221 two-tail test
  - p < .00802111 one-tail test

© 1994 by Gary L. Canivez, Ph.D.  
All rights reserved.
### Discrepancy Agreement Table

WJ-R ACH

<table>
<thead>
<tr>
<th>WIAT</th>
<th>WJ-R ACH</th>
</tr>
</thead>
<tbody>
<tr>
<td>Not Discrepant</td>
<td>Not Discrepant</td>
</tr>
<tr>
<td>23</td>
<td>2</td>
</tr>
<tr>
<td>6</td>
<td>2</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>29</strong></td>
</tr>
</tbody>
</table>

**Results**

Overall Agreement (Hit) Rate = 0.7576

- Observed Agreement Po = 0.7576
- Chance Agreement Pc = 0.6951

Kappa = 0.205

Standard Error of Kappa = 0.159628975

Significance Test for Kappa

- Ho: k = 1.284228005
- Z = 0
- p < 0.19906232 two-tail test
- p < 0.09953116 one-tail test

© 1994 by Gary L. Canivez, Ph.D.
All rights reserved.
**Discrepancy Agreement Table**

WIAT/WE (specific correlations) by W-R ACH/WS

<table>
<thead>
<tr>
<th>WJ-R Ach</th>
<th>Not Discrepant</th>
<th>Discrepant</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>WIAT Not Discrepant</td>
<td>26</td>
<td>1</td>
<td>27</td>
</tr>
<tr>
<td>WIAT Discrepant</td>
<td>5</td>
<td>2</td>
<td>7</td>
</tr>
<tr>
<td>Total</td>
<td>31</td>
<td>3</td>
<td>34</td>
</tr>
</tbody>
</table>

**Results**

Overall Agreement (Hit) Rate = .8235

- Observed Agreement $P_o = .8235$
- Chance Agreement $P_c = .7422$

- Kappa = .3154
- Standard Error of Kappa = .1525621

Significance Test for Kappa
- $H_0: k = 2.067705069$
- $Z = 0$
- $p < .03866763$ two-tail test
- $p < .01933382$ one-tail test

© 1994 by Gary L. Canivez, Ph.D.
All rights reserved.
**Discrepancy Agreement Table**

WIAT/WE (default correlations) by W-R ACH/WS

<table>
<thead>
<tr>
<th>WJ-R Ach</th>
<th>Not Discrepant</th>
<th>Discrepant</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>WIAT Not Discrepant</td>
<td>26</td>
<td>1</td>
<td>27</td>
</tr>
<tr>
<td>WIAT Discrepant</td>
<td>5</td>
<td>2</td>
<td>7</td>
</tr>
</tbody>
</table>

Total | 31 | 3 | 34 |

**Results**

Overall Agreement (Hit) Rate = .8235

Observed Agreement $P_o = .8235$

Chance Agreement $P_c = .7422$

Kappa = .3154

Standard Error of Kappa = .15256261

Significance Test for Kappa $\text{Ho: } k = 2.067705069$

$Z = 0$

$p < .03866763$ two-tail test

$p < .01933382$ one-tail test
### Discrepancy Agreement Table

WIAT/W (discrepancy norms) by W-R ACH/BWL

<table>
<thead>
<tr>
<th></th>
<th>WIAT Not Discrepant</th>
<th>Discrepant</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>WJ-R Ach Not Discrepant</td>
<td>25</td>
<td>3</td>
<td>28</td>
</tr>
<tr>
<td>WJ-R Ach Discrepant</td>
<td>5</td>
<td>3</td>
<td>8</td>
</tr>
<tr>
<td>Total</td>
<td>30</td>
<td>6</td>
<td>36</td>
</tr>
</tbody>
</table>

### Results

Overall Agreement (Hit) Rate = .7778

- Observed Agreement Po = .7778
- Chance Agreement Pc = .6852

Kappa = .2942

- Standard Error of Kappa = .164088703

Significance Test for Kappa

- Ho: $k = 1.792932695$
- $Z = 0$
- $p < .07298358$ two-tail test
- $p < .03649179$ one-tail test

© 1994 by Gary L. Canivez, Ph.D.
All rights reserved.
Discrepancies Between the WISC-III, WIAT, and WJ-R

**Discrepancy Agreement Table**

WIAT/W (specific correlations) by W-R ACH/BWL

<table>
<thead>
<tr>
<th>WIAT Discrepant</th>
<th>WJ-R Ach Not Discrepant</th>
<th>WJ-R Ach Discrepant</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Not Discrepant</td>
<td>29</td>
<td>2</td>
<td>31</td>
</tr>
<tr>
<td>Discrepant</td>
<td>3</td>
<td>2</td>
<td>5</td>
</tr>
<tr>
<td>Total</td>
<td>32</td>
<td>4</td>
<td>36</td>
</tr>
</tbody>
</table>

**Results**

Overall Agreement (Hit) Rate = .8611

- Observed Agreement \( P_0 = .8611 \)
- Chance Agreement \( P_c = .7809 \)

- Kappa = .366
- Standard Error of Kappa = .165509645

Significance Test for Kappa

\[ H_0: \kappa = 2.211351484 \]

\[ Z = \]

\[ p < .02701141 \quad \text{two-tail test} \]

\[ p < .01350571 \quad \text{one-tail test} \]

© 1994 by Gary L. Canivez, Ph.D.
All rights reserved.
Discrepancies Between the WISC-III, WIAT, and WJ-R

Discrepancy Agreement Table

WIAT/W (default correlations) by W-R ACH/BWL

<table>
<thead>
<tr>
<th>WJ-R ACH</th>
<th>Not Discrepant</th>
<th>Discrepant</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>WIAT</td>
<td>29</td>
<td>2</td>
<td>31</td>
</tr>
<tr>
<td>Not Discrepant</td>
<td>3</td>
<td>2</td>
<td>5</td>
</tr>
<tr>
<td>Discrepant</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>32</td>
<td>4</td>
<td>36</td>
</tr>
</tbody>
</table>

Results

Overall Agreement (Hit) Rate = 0.8611

Observed Agreement Po = 0.8611
Chance Agreement Pc = 0.7809

Kappa = 0.366
Standard Error of Kappa = 0.165509645

Significance Test for Kappa
Ho: k = 2.211351484
0  Z =
p < 0.02701141 two-tail test
p < 0.01350571 one-tail test

© 1994 by Gary L. Canivez, Ph.D.
All rights reserved.