Referred Students' Performance on the Reynolds Intellectual Assessment Scales and the Wechsler Intelligence Scale for Children Fourth Edition
Oliver W. Edwards and Rachel V. Paulin
*Journal of Psychoeducational Assessment* 2007; 25; 334 originally published online Jul 19, 2007; DOI: 10.1177/0734282907300453

The online version of this article can be found at:
http://jpa.sagepub.com/cgi/content/abstract/25/4/334

Additional services and information for *Journal of Psychoeducational Assessment* can be found at:

- **Email Alerts:** http://jpa.sagepub.com/cgi/alerts
- **Subscriptions:** http://jpa.sagepub.com/subscriptions
- **Reprints:** http://www.sagepub.com/journalsReprints.nav
- **Permissions:** http://www.sagepub.com/journalsPermissions.nav

**Citations** (this article cites 5 articles hosted on the SAGE Journals Online and HighWire Press platforms):
http://jpa.sagepub.com/cgi/content/refs/25/4/334
Referred Students’ Performance on the Reynolds Intellectual Assessment Scales and the Wechsler Intelligence Scale for Children—Fourth Edition

Oliver W. Edwards
Rachel V. Paulin
University of Central Florida, Orlando

This study investigates the convergent relations of the Reynolds Intellectual Assessment Scales (RIAS) and the Wechsler Intelligence Scale for Children—Fourth Edition (WISC-IV). Data from counterbalanced administrations of each instrument to 48 elementary school students referred for psychoeducational testing were examined. Analysis of the 96 protocols reveals significant positive correlations between conceptually similar RIAS and WISC-IV sub-tests. However, overall mean scores were significantly higher for the RIAS.

Keywords: WISC-IV; RIAS; convergent relations

Examiners often have flexibility regarding which intelligence test they choose to administer when conducting comprehensive evaluations in elementary schools. Some examiners consider briefer measures of intelligence more desirable when the various tests considered for use possess similar psychometric properties (Donders, 1997). When seeking information about an examinee’s cognitive ability that will result in diagnosis or educational decision making, researchers suggest examiners administer the full standard scale of the measure rather than a short form of the test (cf. Edwards & Oakland, 2006; Sattler, 2001).

The most recent iteration of the historic Wechsler scales, the Wechsler Intelligence Scales for Children—Fourth Edition (WISC-IV; Wechsler, 2003), measures the intellectual functioning of children 6 through 16 years, 11 months old. The Reynolds Intellectual Assessment Scales (RIAS), a briefer, more recently developed instrument, were standardized on persons 3 to 94 years old (Reynolds & Kamphaus, 2003). The WISC-IV has 10 core subtests and is said to take approximately 65 to 80 minutes to administer (Wechsler). The four subtests that comprise the RIAS standard battery can be administered by a trained examiner in approximately 20 to 25 minutes (Reynolds & Kamphaus, 2003). Given that both instruments are considered to possess strong psychometric properties (Elliot, 2004; Yeates & Donders, 2005), if the RIAS, which has a shorter administration time, yields

Authors’ Note: We thank the school psychologists and their supervisors who assisted in the data collection. Address correspondence to Oliver W. Edwards, University of Central Florida, Department of Child, Family, and Community Sciences, Orlando, FL 32816-1250. Electronic mail may be sent to owedward@mail.ucf.edu.
similar scores to the more lengthy WISC-IV, examiners who choose to administer the RIAS rather than the WISC-IV may have more time available to provide other important services such as interventions, counseling, or consultation. However, if the two tests do not measure abilities similarly or generate different mean IQs, then the differences may affect the proportion of students who meet eligibility criteria for placement in special education programs (Edwards & Oakland, 2006; Kanaya, Scullin, & Ceci, 2003; Naglieri, Salter, & Rojahn, 2005). The purpose of this study is to investigate the convergent relations of the RIAS and WISC-IV using scores obtained by students referred for psychoeducational evaluation in three public school districts.

Hypotheses

The theoretical underpinnings of the RIAS and the WISC-IV are comparable and include measures of the constructs comprehension-knowledge (Gc) and fluid reasoning (Gf) in determining an examinee’s overall intellectual ability. In addition, both instruments include a measure of working memory (although the RIAS does not incorporate the measure into its computation of an overall intelligence score). Unlike the WISC-IV, the RIAS does not include a processing speed component. It is hypothesized that positive correlations and similar standard scores will be found between subtests measuring related constructs. Correlations are expected to be larger among conceptually similar constructs than conceptually dissimilar constructs. Determining the tests’ correlational relationships will offer examiners information regarding the extent to which the tests measure similar constructs and can be interpreted similarly (Prewett & Matavich, 1994).

Method

Participants

The participants in this study consisted of elementary school students from three mid-size public school districts in a southeastern U.S. state. All students were referred for psychoeducational evaluation by the Student Support Teams from their individual schools. Students were referred most often because of academic problems. Six students were referred to help determine whether they met eligibility criteria for the gifted program. Intelligence test scores from 26 male and 22 female students were used. The sample consisted of 33 Caucasian American (non-Hispanic), 9 African American (non-Hispanic), 4 Hispanic American, and 2 multiracial students. Participants were obtained from 16 elementary schools. The participants were enrolled in kindergarten through fifth grade. The mean age of participants was 8.22 years (SD = 1.42), with a range from 6 to 12.

Instruments

RIAS. The RIAS (Reynolds & Kamphaus, 2003) is designed to measure general intelligence, while “eliminating dependence on motor coordination, visual-motor speed, and
reading skills” (Elliot, 2004, p. 325). The instrument includes measures of crystallized and fluid intelligence, as well as working memory. The first two factors are the most dominant on the RIAS, aside from the general intellectual ability factor (g), on a hierarchy of abilities (Reynolds & Kamphaus, 2003). The test consists of six subtests, four of which comprise the Composite Intelligence Index (CIX). The CIX is defined as a standard score that presents an estimate of an examinee’s general intelligence (g). The scores from the four subtests, which comprise the Verbal Intelligence Index and the Nonverbal Intelligence Index, are used when calculating the CIX (Reynolds & Kamphaus, 2003).

The Verbal Intelligence Index (VIX) is defined as a standard score that presents an estimate of an examinee’s verbal intelligence. This verbal intelligence largely mirrors an examinee’s “crystallized intellectual functions” (Reynolds & Kamphaus, 2003, p. 2).

The Nonverbal Intelligence Index (NIX) is defined as a standard score that represents an estimate of an examinee’s nonverbal intelligence. This nonverbal intelligence largely mirrors an examinee’s “fluid intellectual functions” (Reynolds & Kamphaus, 2003, p. 2).

The Composite Memory Index (CMX) is defined as a standard score that represents a broad estimate of an examinee’s “verbal and nonverbal memory functions” (Reynolds & Kamphaus, 2003, p. 2). The CMX score is not used when calculating an examinee’s CIX.

The test authors report the RIAS has good reliability and is valid for use in the determination of special education eligibility (Reynolds & Kamphaus, 2003; see also Elliot, 2004). Reliability coefficients for the six subtests range from .90 to .95, and reliability for the four RIAS indexes range from coefficients of .94 to .96. Test-retest reliability ranges from .83 to .91, and interscorer reliability ranges from .95 to 1.00 (Reynolds & Kamphaus).

The authors also note that RIAS scores demonstrate good construct validity, and the verbal subtests measure crystallized intelligence, whereas the nonverbal subtests measure spatial ability and reasoning (Reynolds & Kamphaus, 2003; see also Elliot, 2004). The memory subtests are designed to measure working memory. Correlations between RIAS scores and the Wechsler Intelligence Scale for Children–Third Edition (WISC-III; Wechsler, 1991) Full Scale IQ (FSIQ) are between .60 and .78. Correlations between RIAS scores and the Wechsler Adult Intelligence Scale–Third Edition (WAIS-III; Wechsler, 1997) FSIQ are above .70 (Reynolds & Kamphaus).

The WISC-IV is a measure of general intelligence. The instrument yields five composite scores (FSIQ, Verbal Comprehension Index [VCI], Perceptual Reasoning Index [PRI], Working Memory Index [WMI], and Processing Speed Index [PSI]). These indexes provide measures of verbal intelligence, fluid intelligence, working memory, and processing speed. The WISC-IV FSIQ is obtained using the 10 core subtests (Wechsler, 2003).

Subsequent to the publication of the WISC-IV, the test publishers supplied an additional factor called the General Ability Index (GAI) that is less sensitive to working memory and processing speed than the FSIQ. The GAI is a composite score derived from the sum of scaled scores for three VCI and three PRI subtests.

WISC-IV scores are reported to have good reliability and validity for use in the determination of special education eligibility (Wechsler, 2003; Yeates & Donders, 2005). In addition, WISC-IV scores are said to possess good content, response process, and internal structure validity (Wechsler, 2003). Because the WISC-IV’s structure and psychometric
properties are in all likelihood familiar to most readers of this article, they will not be described in detail. The interested reader is referred to the WISC-IV manual for additional information.

**Procedures**

Examiners included 14 state-certified school psychologists and one school psychologist intern. All examiners were trained in the administration of standardized intelligence instruments and in the administration of the RIAS and the WISC-IV. Examiners counterbalanced the instruments’ administrations. Students assessed using the RIAS and the WISC-IV were not assessed previously with the instruments. It was assumed that procedures used during these students’ assessments were the same as those used during the instruments’ standardizations.

**Results**

Scores were obtained and compared for all 48 students. Table 1 contains the descriptive statistics, *t*-test results, and correlations for the two measures. The Bonferroni correction method was employed to decrease the occurrence of Type I errors. A corrected alpha level of .01 was used to determine significant differences. Data found in Table 1 indicate the distribution of scores is generally normal. The sample is composed of students referred for academic problems as well as high academic achievement (i.e., gifted referrals). Given participants such as the aforementioned and with sample sizes greater than 30, violations of normality are unlikely to affect the findings (Tabachnick & Fidell, 1996).

Pearson products were computed for composite pairs. Results of one-tailed tests indicate significant correlations exist between related composite constructs (Table 2). Four correlations are above .86 and are significant at the .00 level. Although significant at the .00 level, NIX-PRI has the lowest correlation among conceptually similar constructs (r = .72). In addition, divergence between conceptually dissimilar constructs was examined. As predicted, correlations are larger among conceptually similar than conceptually dissimilar constructs. Furthermore, because CMX scores are not used in the computation of the CIX and it was not administered for all reported cases, analyses were not conducted with CIX and WMI.

Results of five two-tailed *t* tests indicate significantly higher scores for the RIAS than for the WISC-IV (Table 2). However, effect sizes were small among all cases examining conceptually similar constructs. Generally, 0.2 is considered a small effect, 0.5 a moderate effect, and 0.80 a large effect (Cohen, 1988). Given the range and distribution of scores, caution must be exercised when interpreting these data.

**Discussion**

The results of this study support the hypothesis that a relationship exists between RIAS composite scores and WISC-IV composite scores that seem to measure conceptually similar constructs, namely, comprehension-knowledge, fluid reasoning, and general intelligence. The
strong, significant correlations and substantial shared variances support the convergent validity of the composites. Furthermore, correlations are stronger between the RIAS and WISC-IV than the RIAS and WISC-III.

Table 1
Range, Means, Standard Deviations, and Statistical Significance for Test of Normality

<table>
<thead>
<tr>
<th>Construct</th>
<th>N</th>
<th>Range</th>
<th>Mean</th>
<th>SD</th>
<th>Kolmogorov-Smirnov Test of Normality</th>
</tr>
</thead>
<tbody>
<tr>
<td>CIXa</td>
<td>48</td>
<td>55-135</td>
<td>102.3</td>
<td>15.9</td>
<td>.20</td>
</tr>
<tr>
<td>FSIQc</td>
<td>48</td>
<td>68-126</td>
<td>95.8</td>
<td>17.5</td>
<td>.03</td>
</tr>
<tr>
<td>GAIc</td>
<td>48</td>
<td>64-129</td>
<td>98.7</td>
<td>17.9</td>
<td>.04</td>
</tr>
<tr>
<td>VIXa</td>
<td>48</td>
<td>52-132</td>
<td>101.4</td>
<td>19.6</td>
<td>.16</td>
</tr>
<tr>
<td>VCIc</td>
<td>48</td>
<td>57-130</td>
<td>96.8</td>
<td>17.3</td>
<td>.20</td>
</tr>
<tr>
<td>NIXa</td>
<td>48</td>
<td>74-132</td>
<td>104.7</td>
<td>12.8</td>
<td>.20</td>
</tr>
<tr>
<td>PRIc</td>
<td>48</td>
<td>73-127</td>
<td>99.2</td>
<td>16.3</td>
<td>.02</td>
</tr>
</tbody>
</table>

a. Values above .05 indicate normality. With samples sizes above 30, violations of normality are not considered problematic (Tabachnick & Fidell, 1996).

b. Reynolds Intellectual Assessment Scales (RIAS) tests: CIX = Composite Intelligence Index; VIX = Verbal Intelligence Index; NIX = Nonverbal Intelligence Index.

c. Wechsler Intelligence Scales for Children—Fourth Edition (WISC-IV) tests: Full Scale IQ = FSIQ; VCI = Verbal Comprehension Index; PRI = Perceptual Reasoning Index; General Ability Index = GAI.

Table 2
Correlations, Paired Difference Mean, Paired Difference Standard Deviation, Statistical Significance, and Effect Sizes for Tests and Test Pairs

<table>
<thead>
<tr>
<th>Similar and Dissimilar Test Pairs</th>
<th>Construct Conceptually Similar/ Dissimilar Test Pairs</th>
<th>Correlations Between</th>
<th>Paired Difference</th>
<th>Paired Difference Standard Deviation</th>
<th>t</th>
<th>df</th>
<th>Cohen’s d</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td></td>
<td>Mean</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CIX-FSIQ</td>
<td>48</td>
<td>.90*</td>
<td>6.5**</td>
<td>7.4</td>
<td>6.1</td>
<td>47</td>
<td>.40*</td>
</tr>
<tr>
<td>FSIQ-GAI</td>
<td>48</td>
<td>.95*</td>
<td>–2.9**</td>
<td>5.8</td>
<td>–3.4</td>
<td>47</td>
<td>.16*</td>
</tr>
<tr>
<td>CIX-GAI</td>
<td>48</td>
<td>.90*</td>
<td>3.6**</td>
<td>7.9</td>
<td>3.2</td>
<td>47</td>
<td>.22*</td>
</tr>
<tr>
<td>VIX-VCI</td>
<td>48</td>
<td>.90*</td>
<td>4.6**</td>
<td>8.8</td>
<td>3.6</td>
<td>47</td>
<td>.25*</td>
</tr>
<tr>
<td>VIX-PRI</td>
<td>48</td>
<td>.71*</td>
<td>2.2</td>
<td>14.1</td>
<td>1.1</td>
<td>47</td>
<td>ns</td>
</tr>
<tr>
<td>NIX-PRI</td>
<td>48</td>
<td>.72*</td>
<td>5.5**</td>
<td>11.4</td>
<td>3.3</td>
<td>47</td>
<td>.37*</td>
</tr>
<tr>
<td>NIX-VCI</td>
<td>48</td>
<td>.53*</td>
<td>7.9**</td>
<td>15.2</td>
<td>3.6</td>
<td>47</td>
<td>.51b</td>
</tr>
</tbody>
</table>

Note: CIX = Composite Intelligence Index; FSIQ = Full Scale IQ; GAI = General Ability Index; VIX = Verbal Intelligence Index; VCI = Verbal Comprehension Index; PRI = Perceptual Reasoning Index; NIX = Nonverbal Intelligence Index.


b. Moderate effect size (based on Cohen, 1988).

* Statistically significant at p < .00, one-tailed test. ** Statistically significant at p < .00, two-tailed test.
Although correlations between CIX-FSIQ and between CIX-GAI were statistically significant, t test results reveal RIAS scores are significantly higher. Moreover, despite the similarities between the CIX and GAI in that both reduce requirements for processing speed and working memory, mean scores on the former are significantly higher than the latter. These results suggest that examiners may expect to obtain higher overall IQs when administering the RIAS rather than the WISC-IV. Furthermore, the significant mean difference, coupled with the significant correlation and the effect size of .40 between the CIX and FSIQ, indicates the difference between individual examinee test scores possesses substantial variation around the mean difference. Variance of this nature suggests performance on one test will not reliably predict scores on a subsequent administration of the other test (cf. Prewett & Matavich, 1994).

Limitations associated with this study include the use of students referred for academic difficulties and high achievement as examinees. In addition, a relatively small sample size was obtained from three school districts in a single state. Thus, the noted findings are considered preliminary, and additional studies are needed.

Although available research data suggest that the RIAS and WISC-IV are psychometrically sound intelligence measures, mean score differences between the two tests may have implications for educational decision making and programming. Mean IQ differences between similar intelligence tests may affect the proportion of children classified as special education eligible as a result of which test is administered. For example, the CIX and GAI may affect special education eligibility because of their higher means than the FSIQ. If CIX and GAI scores load lower on g than the FSIQ (as implied by Jensen, 1998), the less these former two scores are predictive of later overall academic achievement in reading, arithmetic, and writing, as well as attainment in any number of areas (Edwards, in press). Future research should investigate the g loadings and predictive validity associated with the CIX and GAI. In addition, future research should determine the degree to which new and revised intelligence tests influence educational decision making as a function of their mean IQ differences and/or construct divergence.

References


