

1-1-2008

Reynolds Intellectual Assessment Scales Racial Bias in Children with Disabilities

Veria Hicks

Follow this and additional works at: <http://mds.marshall.edu/etd>

 Part of the [School Psychology Commons](#)

Recommended Citation

Hicks, Veria, "Reynolds Intellectual Assessment Scales Racial Bias in Children with Disabilities" (2008). *Theses, Dissertations and Capstones*. Paper 641.

This Thesis is brought to you for free and open access by Marshall Digital Scholar. It has been accepted for inclusion in Theses, Dissertations and Capstones by an authorized administrator of Marshall Digital Scholar. For more information, please contact zhangj@marshall.edu.

Reynolds Intellectual Assessment Scales Racial Bias in Children with Disabilities

Thesis submitted to
the Graduate College of
Marshall University

in partial fulfillment of
the requirement for the degree of
Educational Specialist
in School Psychology

by

Veria Hicks

Sandra S. Stroebe, Ph.D, Committee Chairperson
Fred Jay Krieg, Ph.D
Stephen O'Keefe, Ph.D

Marshall University

July 2008

ABSTRACT

Reynolds Intellectual Assessment Scales Racial Bias in Children with Disabilities

By Veria Hicks

Much has been written regarding bias in intellectual tests, in particular racial bias (Naglieri & Rojhan, 2001). Depending on the way intelligence is measured, a difference may be found between racial and ethnic groups. In the second half of the 20th century on the average, the mean IQ score of African Americans was one standard deviation below that of Caucasian students (Gerig and Zimbardo, 2002). This study addressed whether racial bias existed in the Reynolds Intellectual Assessment Scales (RIAS), and examined the relationship of the RIAS to other intelligence tests. The results were analyzed using a simple linear regression. There was no significant difference between the scores of the African American students and the Caucasian students. There was a moderate correlation between the RIAS and other intelligence tests.

Dedication

This is dedicated to my grandsons, Kaleb, Jacob, Chance, and Grayson, whom I love very much and who are the reason I am who I am.

Acknowledgement

I would like to thank Dr. Stroebel for all of her hard work, input and patience with me as I completed this study. I would also like to express my gratitude to Dr. O'Keefe for his help with the statistical aspects, and Dr. Krieg for his suggestions to make the research stronger. I would also like to thank my family for their encouragement along the way. Last but not least, I want to thank God for allowing me to follow through with the dream of being a school psychologist.

TABLE OF CONTENTS

ABSTRACT.....ii

DEDICATION.....iii

ACKNOWLEDGEMENT.....iv

TABLE OF CONTENTS.....v

LIST OF TABLES.....vi

CHAPTER I: Literature Review.....1

CHAPTER II: Method..... 8

CHAPTER III: Results.....10

CHAPTER IV: Discussion.....11

REFERENCES.....16

List of Tables

Comparison of Reliability and Validity of Intelligence Tests.....	13
Analysis of Linear Regression.....	14
Pearson Correlation.....	15

Chapter I

As school psychologists, we are ethically and legally required to select test instruments that are not racially or culturally biased. When new tests are developed they need to be assessed for bias. One recently developed test is the Reynolds Intellectual Assessment Scales (RIAS). This test is considered a short form intelligence test as it takes approximately 25 minutes to administer (Reynolds & Kamphaus, 2003). This study proposes to assess the RIAS in the area of racial bias and its relationship to other longer IQ tests.

Review of Literature

History

Intelligence testing has been practiced for a long time. According to Gerrig & Zimbardo (2002) assessment techniques were commonly used in China for civil service testing over 4,000 years ago. Individuals who wanted to be employed by the government had to pass three phases of the exam to be hired, (p.291). In the Western world, Sir Francis Galton was the first to state there were four important ideas about the assessment of intelligence (Gerrig & Zimbardo, 2002). Sir Francis Galton believed differences in intelligence were quantifiable in terms of degrees of intelligence, differences among individuals formed a bell shaped curve, intelligence could be measured by objective tests and these tests could be correlated (Gerrig & Zimbardo, 2002). In 1905 Alfred Binet created an intelligence test, at the behest of the French Government, to separate mentally impaired children from normal children (Feldman, 2001). There were four important features of Binet's test. The first was the ability to interpret the scores as an estimate of current performance. Next, he wanted the scores to be used as a means to identify

children who needed help. Thirdly, he tried to identify areas of performance in which special education could help those children identified as mentally impaired. Fourthly, he collected data to see if his test was valid (Gerrig & Zimbardo, 2002).

Binet had a great impact on the United States with men such as Lewis Terman, Edward Thorndyke, and Robert Yates. Lewis Terman adapted Binet's test questions for use with American school children. He revised the Binet and called the test the Stanford-Binet (SB) in 1916 since he was teaching at Stanford University (Gerrig & Zimbardo, 2002; Feldman, 2001).

Others were also independently developing intelligence tests. David Wechsler developed an intelligence test for adults in 1939, which he published as the Wechsler Bellevue Intelligence Scale. After some changes the test was published as the Wechsler Adult Intelligence Scale (WAIS) in 1955 (Gerrig & Zimbardo, 2002; Feldman, 2001; Nagleeri & Rojhan, 2001). Wechsler's tests are older as he adopted many of the methods used by the U.S. military in the early part of the 1900's. Lewis Terman, Edward Thorndyke, and Robert Yates also developed intelligence tests to assist in determining who would be leaders for the U.S. military during World War I (Gerrig & Zimbardo, 2002). Wechsler developed the Wechsler Scale for Children (WISC) in 1949 and Primary and Preschool Intelligence Scale in 1967 (Reynolds & Gutkin, 1999). For a long period of time the WISC and SB dominated the field of intellectual assessment of children.

The Kaufman Assessment Battery for Children (K-ABC) was developed in 1983 by Kaufman and Kaufman and is another assessment instrument used to determine the intelligence of children. However, this scale is "predicated on the distinction between

problem solving and knowledge of facts”, (Reynolds & Gutkin, 1999). According to Reynolds and Gutkin (1999), this test was different from other intelligence tests in which a person’s acquired information and applied skills influence the IQ score obtained.

Many revisions have since been completed on several of these tests. The Wechsler Intelligence Scale for Children –Fourth Edition (WISC-IV) and the Stanford-Binet – Fifth Edition (SB-V) were both revised in 2003. The K-ABC-Second Edition was revised in 2004. According to Minton and Pratt (2006) the norming process for the revised tests such as the SB-V and WISC-IV posed problems with identifying gifted students. “ Both scales reflect changes that go beyond simple item updates and new norms; there are refinements in the conceptualization of intelligence and corresponding changes in item subject matter” (p.233). Because of these changes, use of these new tests may influence the number and type of students that are deemed to be eligible for services particularly gifted children, who scored lower on the new revised tests (Minton & Pratt, 2006). Research is needed to see the impact of the changes on the identification of children with handicapping conditions and children from different ethnic backgrounds.

More options are available to practitioners than ever before. Along with the WISC-IV, SB-V, K-ABC-II, there is the Woodcock-Johnson Cognitive (WJ-COG) and the Cognitive Assessment System (CAS). While abbreviated procedures were available for many of the intelligence tests, these were rarely recommended for use by school psychologists, especially when used for educational decisions (Sattler and Dumont, 2004). Reynolds developed a shorter intelligence test, The Reynolds Intellectual Assessment Scales (RIAS) that he advocated as highly correlated with currently used instruments yet with a shorter administration time. He stated that if the full scale scores

are the same, why not use the test with a shorter administration time (Personal Communication, May 18, 2006). With all the new intelligence tests on the market to choose from, research is needed to guide practitioners about which ones to use.

Laws

Not only does our ethical code require school psychologists to select instruments that are appropriate, but we are mandated by law to utilize appropriate instruments when evaluating children for special education. In 1975, the Congress of the United States passed into legislation, Public Law 94-142. This law required states to provide special education services to children with special needs (Willis, D. 1977). In order for children to be identified as needing special education services, a full psychological evaluation was required. Scores on the assessments were used to determine eligibility.

The No Child Left Behind Act of 2001 reauthorized the Elementary and Secondary Education act of 1965. This act provided funding for states to implement reading readiness to preschool age children. IDEA, 2004 and subsequent Regulations in August 2006, are the most recent laws which ensure students receive a free and appropriate public education. Under IDEA, students who are referred for special education services must be evaluated in order to determine eligibility. These evaluations consist of not only achievement tests but also intelligence tests (U.S. Department of Education, 2007). These tests must be nondiscriminatory, valid and reliable (U.S. Department of Education, 2007). In the NASP Code of Ethics, the code on test instruments states that school psychologists need to use appropriate assessment procedures when evaluating students and the APA code of ethics states psychologists must use assessment instruments in which the validity and reliability have been established” (Jacob & Hartshorne, 2007).

Test Bias

Intelligence test bias has been the center of controversy for years by many people in the educational, psychological, and legal fields. Bias has been identified by investigators choosing to explain the score differences, differential factorial structure appropriateness of test content and differences in regression lines (Naglieri & Hill, 1986). In the study by Naglieri & Hill (1986), regression lines of the WISC-R and the K-ABC were examined. The sample included only African American and Caucasian Children. The results suggested the regression lines for the African American and the Caucasian students indicated no significant difference. In a study to determine eligibility for special education services Kush, Watkins, Ward, Ward, Canivez, & Worrell (2001) evaluated African American and Caucasian students. The results suggested a 15 point difference between African American and the Caucasian students. The African American students exhibited slightly lower Full Scale IQ scores than did the Caucasian students. In a comparative study of the WISC-III and the Cognitive Assessment System (CAS) Naglieri & Rojhan (2001) identify disproportionately more African American children as having mental impairment on the WISC-III than on the CAS. This finding appears to be consistent with earlier criticism of the WISC-III as being responsible for the over representation of African American children placed in special education programs (Naglieri & Rojhan, 2001).

Saccuzzo, Johnson & Russell (1992) conducted a study where Verbal IQ versus Performance IQ scores among gifted African American, Caucasian, Filipino, and Hispanic children were studied. Results of this study indicated significant differences occurred among the different populations. The African American and Caucasian students

obtained significantly higher Verbal IQ scores than Performance IQ scores. The Hispanic students had a higher Verbal IQ score but the difference was not significant. The Filipino students scored higher on the Performance IQ than the Verbal IQ score. The differences in Verbal IQ – Performance IQ pattern appears to have been dependent upon ethnic background, the interaction of ethnicity and the size of the student’s Verbal IQ-Performance IQ discrepancy.

The above studies suggest a need to evaluate new and revised IQ tests for differences between racial groups. In the RIAS, the authors excluded gender and ethnicity differential items in order to eliminate any bias in the assessment (Reynolds & Kamphaus, 2003).

Abbreviated Tests

As well as assessing the presence of racial bias in test instruments, the utility of abbreviated intelligence forms needs to be examined. Donders (2001) wanted to find a more time efficient test for use in practice and conducted research using the abbreviated form of the WISC-III. In an initial study in 1997, Donders found no significant differences between the WISC-III and WISC-III SF IQ scores. Yet as he continued to use the WISC-III Short Form (SF) for neuropsychological evaluations from 1998 through 2000, significant differences between scores on the long form and short form were found. Minshew, Turner, and Goldstein (2005) conducted a study with high functioning adults and children diagnosed with Autism, using the short form of the Wechsler Intelligence scales. Their study questioned whether the predictive values of the short forms were comparable to the long forms. Results showed short forms were as good a predictor as long forms for both adults and children with high functioning Autism.

As seen in the above studies there have been varied findings regarding whether abbreviated forms of IQ tests were as good as the long form tests. Even though the research data is unclear about the validity of using abbreviated tests, Jerome Sattler and Ron Dumont (2004) noted in their book *Assessment of Children* that short form IQ tests may be used only for screening purposes. When determining the child's IQ, they reported short form IQ tests magnify the effect of any administrative errors and gives too much weight to each subtest. Their recommendation was to forgo short form tests for determining programming needs of children.

In the NASP Code of Ethics there is no mandate as to whether short form or long form IQ tests are to be utilized. What is required is that the tests used are appropriate and data based (Jacob & Hartshorne, 2007). After reviewing the research it appears there is still no consensus as to whether IQ scores obtained from short form IQ assessments are comparable to the long forms. While the majority of studies did not find significant differences when using the short form versus the long form, many caution the use of short form for decision making (Sattler & Dumont, 2004).

Reliability and Validity

When considering the technical adequacy of tests, it is important to consider the reliability and validity of a test. Please refer to Table 1 for a review of the reliability and validity of some of the most widely used intellectual assessments. The WISC-IV reliability ranges from .86 to .97 while the validity is reported as .83 (Sattler and Dumont, 2004). The reliability for the Stanford-Binet -V ranges from .84 to .98. Internal reliability ranges from .95 to .98 and the validity was reported as strong (Roid and Barram, 2004). Reliability for the K-ABC-II ranges from .84 to .97 with accompanying

strong (.70's to .80's) validity scores (Flanagan and Kaufman, 2004). The CAS has a reliability of .96 with a strong validity being reported (Pearson Assessment group, 2004). Looking at the reliability for the WJ-III Cognitive the reliability is .97 with a strong validity being reported (WJ-III Test of Cognitive Abilities, 2004). The RIAS has a reliability of .83 to .91 with a strong validity reported (Reynolds and Kamphaus, 2003).

As seen by Table 1 most of the tests discussed appear to have good reliability and validity

With all the available cognitive assessment instruments for children, research is needed to ensure that the instruments are non-discriminatory. This study attempts to assess racial bias in the RIAS as well as how well it is correlated with other intelligence measures.

Ho₁: There is no significant difference in IQ scores of African American students compared to Caucasian students in the RIAS.

Ho₂: The RIAS is highly correlated with other IQ tests

Chapter II

Method

Subjects

Ninety-four students in the Kanawha County School system were given the Reynolds Intellectual Assessment Scales (RIAS), during an 18 month period of time. The students had previously been given other IQ tests such as the (CAS), Differential Ability Scale (DAS), Stanford-Binet Fifth Edition (SB-V), Woodcock-Johnson Test of Cognitive Abilities-III (WJ-Cognitive), Kaufman Assessment Battery for Children-II (K-ABC-II),

and the Wechsler Intelligence Scale for Children-IV (WISC-IV), for either the initial evaluation to determine eligibility or as an annual evaluation to continue receiving special education services. Of the 92 students evaluated, 20 were African-American and 74 were Caucasian. All students were referred for special education evaluation. Students who had an I.Q. score of 50 or below or an IQ of 130 and above were excluded from the study.

Procedure

Data was collected over an 18 month period, from January 2006 to June 2007. The RIAS was given according to the instructions from the test manual. A Certified School Psychologist administered the RIAS as either part of the student's triennial re-evaluation or for an initial evaluation to determine eligibility for services in special education.

Instrument

The Reynolds Intellectual Assessment Scales (RIAS) is a short form intelligence test which can be administered across all developmental stages from ages three up through age 94. The RIAS includes two- subtest Verbal Intelligence Index (VIX), a two –subject Nonverbal Intelligence Test (NIX), and a Composite Intelligence Index (CIX). There are four subtests. The total raw score for each subtest is transferred to the RIAS Score Summary Table. Each subtest raw score is converted to T scores then the T scores are converted to index scores. A conversion table is included in the test manual for the T scores.

Chapter III

Results

Analysis of Data

A simple linear regression was calculated predicting racial bias in the RIAS. The regression equation was not significant ($F(3,90)= 50.241, p >.05$) with an R squared of .626. Race cannot be used to predict scores on the RIAS (Table 1). There was no significant difference between the IQ scores of African-American and Caucasian students who were assessed using the RIAS. An analysis did show that other IQ tests are more of a predictor than race or gender for the RIAS (See Table 2 and 3).

A Pearson Product Moment Correlation was used to see if there was a significant difference between the RIAS and other intellectual tests. Consistent with the linear regression, the results suggested there was a moderate correlation between the RIAS and other IQ tests at the 0.01 level (See Table 3).

Chapter IV

Discussion

School Psychologists are mandated by law and ethical codes to select test instruments that are racially unbiased and valid. While previous instruments were found to exhibit differences in racial groups, current instruments claim to be unbiased (Naglieri & Hill, 1986; Reynolds & Kamphaus, 2003; Sattler & Durmont, 2004). Research studies are needed to examine the claims by test publishers that their instruments are unbiased.

This study attempted to examine the RIAS for racial bias. Consistent with the author's claims, the RIAS was not biased based on race. More studies are needed with larger and more diverse groups including non special education groups to confirm the findings in this study.

This study also examined the relationship between the RIAS and other IQ tests. A moderate correlation was found ($r=.78, p>0.01$). While this differs from the high correlation reported by Reynolds (Personal Communication, May 18, 2006), this study may have slightly different results due to the smaller sample size and limited population. Only children in or referred for special education were included in the study.

As the RIAS can be administered in as little as 25-30 minutes, is a complete intelligence test, and not considered a short form of an intelligence test, it may be prudent for school systems to administer to save time when doing evaluations. The school district from which this data was obtained is currently using the RIAS for re-evaluations to speed up assessments.

There are some limitations to this study which need to be addressed. The first is the number of African American students. Compared to the number of Caucasian students,

there was a small number of African American students. What would have been ideal would have been to have an equal number for both groups. Another area of concern was the lack of information concerning the SES of the students. Previous studies have found differences between economic groups on intelligence tests. Some suggesting that economics rather than race is the most important issue (Flanagan & Kaufman, 2004; Grados, & Russo-Garcia, 1999 ; Roid & Barram, 2004; Sattler & Dumont, 2004; Turkheimer, Haley, Waldron, D'Onofrio, & Gottesman, 2003). Another limitation is that the study only included children that were referred for special education thus the findings are not applicable to the general population.

Table 1

COMPARISON OF RELIABILITY AND VALIDITY OF DIFFERENT
INTELLIGENCE TESTS

	Test Retest Reliability	Inter-rater Reliability	Validity(Correlations with other measures)
WISC-IV	.88 to .97	..86 to .93	.83
SB-V	.95 to .98	.84 to .89	.90
K-ABC-II	.84 to .97	.85	Strong (.70's and .80's
CAS	.96	.87	Strong
WJ-COG	.97	.71 to .76	Strong
RIAS	.83 to .91	.95 to 1.00	Strong

Table 2

ANALYSIS OF LINEAR REGRESSION

Coefficients

Model	Unstandardized Coefficients		Standardized Coefficients Beta	t	Sig.
	B	Std. Error			
1 (Constant)	23.189	7.632		3.038	.003
Race	3.384	2.261	.097	1.497	.138
Sex	-4.252	2.045	-.134	-2.079	.040
Other IQ	.773	.065	.768	11.911	.000

Dependent Variable: RIAS

Table 3

PEARSON CORRELATION

			RIAS	Other IQ Tests
RIAS	Pearson Correlation	1		.777
	Sig. (1- tailed)			.000
	N	91		91
Other IQ Tests	Pearson Correlation	.777		1
	Sig. (1-tailed)	.000		
	N	91		91

References

- Donders, J. (1997). A Short Form of the WISC-III for Clinical Use. *Psychological Assessment, 9*(1), 15-20.
- Donders, J. (2001). Using a short Form of the WISC-III: Sinful or Smart. *Child Neuropsychology, 7*(2), 99-103.
- Feldman, R. S. (2001). *Child Development*. Upper Saddle River, New Jersey: Prentis Hall.
- Flanagan, D. P. & Kaufman, A. S. (2004). *Essentials of WISC-IV Assessment*. Hoboken, New Jersey, USA: Jon Wiley & Sons.
- Gerrig, R. J., & Zimbardo, P. G. (2002). *Psychology and Life*. Boston: Allyn and Bacon A. Pearson Education Company.
- Grados, J. J., Russo-Garcia, K.A. (1999). Comparison of the Kaufman Brief Intelligence Test and the Wechsler Intelligence Scale for Children-Third Edition in Economically Disadvantaged African American Youth. *Journal of Clinical Psychology, 55*(9), 1063- 1071.
- U. S. Department of Education. (2007). *Twenty-third annual report to congress on the implementation of the Individuals with Disabilities Education Act*. Washington, DC: Author. Retrieved November 14, 2007, from <http://www.ed.gov/about/reports/annual/osep/2001/section-i.pdf>
- Jacob, S. & Hartshorne, T.S. (2007). *Ethics and Law*. Hoboken, New Jersey, USA: Jon Wiley & Sons.

- Kaufman, A. S., Kaufman, J. C., Balgopal, B., & McClean, J. E. (1996). Comparison of Three WISC-III Short Form : Weighing Psychometric, Clinical, And Practical Factors. *Journal of Clinical Child Psychology, 25*(1), 97-105.
- Kush, J. C., Watkins, M. W., Ward, T. J., Ward S. B., Canivez., G.L., & Worrell, F. C. (2001). Construct Validity of the WISC-III foe White and Black Students from the WISC-III Standardization Sample and for Black Students Referred for Psychological Evaluation. *School Psychology Review, 30*(1), 70-85.
- Minsheu, N. J., Turner, C. A., & Goldstein, G. (2005). The Application of Short Forms of the Wechsler Intelligence Scales in Adults and Children with High Functioning Autism. *Journal of Autism and Developmental Disorders, 35*(1), 52.
- Minton, B. & Pratt, S (2006). Identification Discrepancies. *Roeper Review, 28*(4), 232-236.
- Naglieri, J. A., & Hill, D. S. (1986). Comparison of WICS-R and K-ABC regression Lines for Academic Prediction with Black and White Children. *Journal of Clinical Child Psychology, 15*(4), 352-355.
- Naglieri, J., & Rojhn, J. (2001). Intellectual Classification of Black and White Children in Special Education Programs Using the WISC-III and the Cognitive Assessment System. *American Journal on Mental Retardation, 106*(4), 359-367.
- Pearson Assessment group Web Site, (2007). *Reliability of the RIAS*. Retrieved January 15, 2008 from [http; // www.pearsonassessment.com](http://www.pearsonassessment.com).
- Reynolds, C. R., & Gutkin, T. (1999). *The Handbook of School Psychologist*. Hoboken, New Jersey: John Wiley & Sons.

Reynolds, C.R., & Kamphaus, R.W.(2003) Reynolds Intellectual Assessment Scales (RIAS) Manual

Roid,G., Barram, R. (2004). *Essentials of Stanford-Binet Intelligence Scales (SB5) Assessment*. Hoboken, New Jersey, USA: John Wiley & Sons.

Saccuzzo, D. P., Johnson, N. E., & Russell, G. (1992). Verbal Versus Performance IQs for Gifted African-American, Caucasian, Filipino, and Hispanic Children. *Psychological Assessment, 4*(2), 239-244.

Sattler, J. M., & Dumont, Ron. (2004). *Assessment of Children*. San Diego, California, USA: Jerome M. Sattler, Publisher, Inc.

Schrank, F.A., Flanagan, D.P., Woodcock, R.W., & Mascolo, J.T.(2004) *Essentials of WJ-III Cognitive Abilities Assessment*, Hoboken, New Jersey, USA: Jon Wiley & Sons

Silverstein, A. B. (1965). 'Validity" of Short Form Intelligence Tests. *Journal of Consulting Psychology, 29*(4), 392-393.

Turkheimer, E., Haley, A., Waldron, M., D'Onofrio B., & Gottesman, I.(2003). Socioeconomic Status Modifies Heritability of IQ In Young Children. *Psychological Science, 14*(6), 623-628.

Valencia, R. R., Rankin, R., & Oakland, T. (1997). WISC-R Factor Structures Among White, Mexican American, and African American Children: A Research Note. *Psychology in the Schools, 34*(1), 11-15.

West Virginia State Board of Education (2007). Public Law 2419 Regulation for the Education of Exceptional Children. Retrieved July 15, 2008, <http://wvde.state.wv.us/policies/>

Willis, D. (1977). Public Law 94-142 and The Rights of Children. *Journal of Clinical Child Psychology*, 6(3), 2.