1-1-2005

The construct validity of the Wechsler Abbreviated Scale of Intelligence (WASI) and the Wide Range Intelligence Test (WRIT): Replication and extension

Gregory S. Wilson
Eastern Illinois University

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 for the purpose of copying it for inclusion in that institution's library or research holdings.

______________________ 5/16/05

Author's Signature Date

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

____

____

____

______________________ ______________________

Author's Signature Date

This form must be submitted in duplicate.

The Construct Validity of the
Wechsler Abbreviated Scale of Intellilgence (WASI) and the
Wide Range Intelligence Test (WRIT) : Replication and Extension

BY

Gregory S. Wilson

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

2005
YEAR

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

5/11/05
Date

Thesis Director

5/11/05
Date

Department/School Head

The Construct Validity of the

Wechsler Abbreviated Scale of Intelligence (WASI) and the

Wide Range Intelligence Test (WRIT): Replication and Extension

Gregory S. Wilson

Eastern Illinois University

Table of Contents

List of Tables and Figure

Acknowledgments

I would like to thank my wife Spring for supporting my efforts through graduate school and during the completion of my thesis. She stayed home with our two boys for three years, to ensure that I would accomplish my goals. Spring, you are a wonderful mother and wife. I would also like to thank my thesis chairperson, Dr. Gary Canivez. I was persistent at times. I appreciate your understanding, as well as the time, effort, and support you have provided me. Thank you.

I would like to thank my other thesis committee members, Dr. Assege HaileMariam and Dr. Michael Havey for their help, as well. Alesia Grigg, my internship supervisor, also played an instrumental role in the completion of my thesis. I would like to thank her for helping me get participants and convincing administrators to let me use their students. Finally, the project would not have been a success if not for the cooperation of the parents, students, and other individuals who agreed to participate in the study.

Abstract

The purpose of this study was to further verify the relationship between the Wechsler Abbreviated Scale of Intelligence (WASI) and the Wide Range Intelligence Test (WRIT) by replicating Collins' (2002) study regarding the convergent validity between these two tests. To assess the convergent validity of the WASI and WRIT, two newer brief intelligence tests, 86 participants from central Illinois public schools were administered both measures in counterbalanced order. All correlations between the WASI and WRIT were significant ($p < .05$). Furthermore, the correlations between the corresponding IQ scales and subtests of the WASI and WRIT were moderate to high, ranging from .55 ($r^2 = .30$) to .82 ($r^2 = .67$). The results provided convergent evidence for the construct validity of the WASI and WRIT and supported the contention made by Collins (2002) that the WASI and WRIT are measuring similar constructs. The results of the joint exploratory factor analysis supported the retention of only one factor (general intelligence "g"). A forced, two-factor solution resulted in 10.50% more variation among WASI/WRIT subtest scores. However, the two factors correlated at a high magnitude ($r = .75$), and all subtests cross-loaded with their theoretical dissimilar factor. The results of the factor analysis supported a conclusion that the WASI and WRIT are best conceptualized as providing strong measurements of general intelligence.

Introduction

A Brief History of Intelligence and Testing

In 1905, Binet and Simon brought out an empirically developed device that was able to differentiate between those who were mentally retarded and students who possessed the ability to learn. This device was viewed as the ultimate operational definition of intelligence during that time. It was eventually revised and adapted for use in the United States and became the Stanford-Binet intelligence test (Thorndike, 1990). A major objective of the work of Binet was to classify by ability level. A current function of intelligence tests remains to make the same basic distinction that concerned educators 100 years ago.

Although some practitioners use the Stanford-Binet revisions in assessment, the Wechsler Scales are currently the most commonly used tests of intelligence (Stinnett, Havey, Oehler-Stinnett, 1994). Wechsler felt that the Binet scales were too verbally loaded for use with adults. Therefore, he designed an instrument with subtests to measure both verbal and nonverbal abilities (Thorndike, 1997). The original test, the Wechsler-Bellevue (Wechsler, 1939), was quite successful. Since the original test, different versions of the instrument have been developed. Each version focuses specifically on different age groups and includes the Wechsler Preschool and Primary Scale of Intelligence-Third Edition (WPPSI-III; Wechsler, 2002), the Wechsler Intelligence Scale for Children-Fourth Edition (WISC-IV; Wechsler, 2003), and the Wechsler Adult Intelligence Scale-Third Edition (WAIS-III; Wechsler, 1997). All of these scales still show a marked similarity to the original 1939 scale (Thorndike, 1997).

Even before modern times, a characteristic of intelligence testing that drew criticism was the lack of a sound theoretical base for the tests (Thorndike, 1990). Neither the original Binet and Simon scale nor the original Wechsler scales were based on theory. They were simply a group of tasks that were verbally and non-verbally divided. Most tests developed within the past 15 to 20 years have been theory driven or at least based on a preexisting theory of intelligence (Ittenbach, Esters, & Wainer, 1997). Factor analytic research conducted using past and current theories have stressed one of two models. One model is the hierarchical one-factor theory "g" advocated by such researchers as Jensen (1980) and Spearman (1927). The other model is a multifactored theory proposed by such investigators as Horn (1985), Gardner (1983), and Thorndike (1927). Even with these advances, the basic means for determining a child's ability remains little different than it was during Simon and Binet's time.

Rational for Using Brief Intelligence Tests: Problems with Short Forms

Over the years, the workloads of many school psychologists have grown, which has resulted in an increased amount of time spent on evaluations. Comprehensive intelligence tests take over two hours to administer, score, and interpret. Due to these circumstances, attempts have been made to decrease the amount of time it takes to complete an evaluation. The most popular solution has been the use of short forms of comprehensive intelligence tests (Kaufman & Kaufman, 2001). Some of the most popular short forms have been abbreviated versions of the Wechsler scales. These measures can achieve their brevity in a number of ways, including the use of two, three, four, and seven subtests or even taking out every second or third subtest item (Kaufman & Kaufman, 2001). Compared to preexisting brief measures with questionable psychometric properties and homogeneous tests which were not even an adequate measure of the intelligence construct, short forms were the best choice. Short forms offered clinicians a reliable and valid assessment option and decreased the time needed to do an evaluation, but today short forms are no longer the superior brief measure of intelligence (Kaufman & Kaufman, 2001).

Short forms are not specifically developed as brief tests. Standardization and technical characteristics are based on the administration of an entire test, not on a subset of subtests. Therefore, the reliability, validity, and norms of short forms are not actually known. Thompson (1987) noted this problem. A study was conducted where one group was given the standard WAIS-R, and another group was given the Vocabulary and Block Design subtests first. In the standard version these subtests were given fifth and sixth, respectively. Results indicated that subjects performed uncharacteristically well when administered the two subtests first. Therefore, norms for short forms are not necessarily valid when derived from norms for the complete battery.

Other problems with short forms have also been addressed. Silverstein (1990) suggested that short form correlations with VIQ, PIQ, FSIQ, VCI, and POI scores would be spuriously high, due to the fact that they are included in calculating the respective IQ and Index scores. However, newer brief tests of intelligence like the Kaufman Brief Intelligence Test (K-BIT; Kaufman & Kaufman, 1990), Wechsler Abbreviated Scale of Intelligence (WASI; The Psychological Corporation, 1999), and Wide Range Intelligence Test (WRIT; Glutting, Adams, & Sheslow, 2000) were specifically developed as brief tests, normed as brief tests, and validated as brief tests. The K-BIT, WASI, and WRIT have excellent psychometric qualities. After the publication of these brief intelligence tests there was no longer a need to create or use short forms in order to decrease the time needed to do an evaluation (Kaufman & Kaufman, 2001).

Rational for Using Brief Intelligence Tests: Problems with Profile Analysis and Ipsative Interpretation

Profile analysis is the interpretation of subtest patterns. Unfortunately, some individuals over interpret these patterns. It is assumed by some that groups of similarly diagnosed individuals represented meaningful categories. However, in regard to the WISC-III, clinically unique profiles are rare in both normal and special education populations (Glutting, McDermott, Watkins, Kush, and Konold, 1997).

In ipsative interpretation test subjects are used as their own norm. This method is intuitively appealing because the resulting profile appears to isolate and magnify aspects of differential ability (McDermott, Fantuzzo, & Glutting, 1990). However, the technique is fraught with problems. One problem is that deviation scores must always sum to zero. Therefore, when performance improves in one area, it appears to deteriorate in another area (McDermott et al., 1990). Another problem is that difference scores have poor reliability. The standard error of difference between two subtest scores is greater than the standard error of measurement of the two scores (Kamphaus, Petoskey, & Morgan, 1997). In addition, ipsatization removes all common variance associated with general intelligence. Consequently, this results in the loss of over 50% of a test's reliable variance (McDermott, Fantuzzo, Glutting, & Watkins, 1992). Furthermore, evidence regarding the effectiveness of ipsative scores for prediction is lacking (McDermott et al., 1990). Lastly, subtests provide little differential information about a child's cognitive abilities. In fact, few WISC subtests can attribute one third or more of their variance to subtest specific variance (Cohen, 1959; Kaufman, 1979; Kamphaus, 1993).

Rationale for Using Brief Intelligence Tests: Closing Remarks

Many would agree that longer tests measure general intelligence better than brief tests, but how much better? General intelligence is a collection of all of an individual's cognitive skills. Even with comprehensive tests, everything can not be measured. Therefore, when school psychologists use comprehensive measures over brief measures, they may not be gaining much regarding their assessment of general intelligence (Glutting et al., 2000). Furthermore, IQ testing is only one part of the assessment process. In other words, an individual's intelligence offers only one piece of information that is needed to make a determination. Effective assessment involves collecting information from multiple sources of data in order to consider context, creating hypotheses that can be confirmed by more than one source of data, making the most parsimonius conclusions, and supporting those conclusions with empirically based research and theory. Therefore, an IQ score is only one part of one step in the whole process.

When doing a psychological assessment, many pieces of information are needed. However, intelligence is only one piece of information. Does the additional administration time of comprehensive IQ tests increase meaningful contributions to differential diagnosis and treatment planning, or are brief tests sufficient? In other words, what are the benefits to using a more comprehensive test? One possible benefit is the availability of subtest interpretation. However, as mentioned earlier, the practices of profile analysis, ipsative interpretation, and single-subtest comparisons are misuses of comprehensive intelligence tests and do not enhance predictive or treatment validity. In fact, neither brief nor comprehensive IQ tests have much utility beyond their ability to measure general intelligence or the fluid/crystallized components of intelligence. In general, brief IQ tests are a quick determination of general intelligence, and, as mentioned earlier, general intelligence is a good predictor of many variables central to education.

Newer Brief Measures of Intelligence

Besides the poorly constructed Slosson Intelligence Test (Jensen & Armstrong, 1985), brief intelligence tests were not an option for practitioners until the development of the Kaufman Brief Intelligence Test (K-BIT; Kaufman & Kaufman, 1990). The K-BIT was designed and normed for brief administration, and it was the only option for nearly a decade. The K-BIT is composed of two subtests and measures an individual's verbal/crystallized and nonverbal/fluid intelligence. The K-BIT has good reliability and validity (Kaufman & Kaufman, 1990).

Wechsler Abbreviated Scale of Intelligence (WASI)

Since the publication of the K-BIT, there have been other tests developed specifically for the purpose of brief administration. One is the Wechsler Abbreviated Scale of Intelligence (WASI; The Psychological Corporation, 1999). The WASI is an individually administered intelligence test and can be used with individuals between the ages of 6 and 89 years. It was standardized nationally, and yields the customary Verbal (VIQ), Performance (PIQ), and Full Scale IQ (FSIQ-4; FSIQ-2) scores. Furthermore, the WASI was linked to the Wechsler Intelligence Scale for Children-Third Edition (WISC-III; Wechsler, 1991) and the Wechsler Adult Intelligence Scale-Third Edition (WAIS-III;

Wechsler, 1997). It provides tables for estimating score ranges on these more comprehensive tests (The Psychological Corporation, 1999).

A major advantage of the WASI is that administration is similar to other Wechsler scales, and most psychologists are familiar with these measures. The WASI is comprised of four subtests (Vocabulary, Block Design, Similarities, and Matrix Reasoning). The design of these subtests is similar to the subtests in previous Wechsler scales. These subtests were specifically chosen because of their strong association with general intelligence. They were also chosen because of their relationship to the constructs of fluid and crystallized intelligence, the theoretical base of the WASI. Administration of all four subtests takes approximately 30 minutes. However, when time is a major constraint, only two WASI subtests (Vocabulary and Matrix Reasoning) are needed for estimating general cognitive functioning. The two test format takes 15 minutes or less to administer (The Psychological Corporation, 1999).

The technical data in the WASI manual shows that standardization was based on a sample of 2,245 individuals between 6 and 89 years of age. The WASI's internal consistency reliability was established via the split-half method. The average reliability coefficients for the entire child sample were .93 for VIQ, .94 for PIQ, .96 for FSIQ-4, and .93 for FSIQ-2. The average reliability coefficients for the adult sample were as follows: .96 for VIQ, .96 for PIQ, .98 for FSIQ-4, and .96 for FSIQ-2. All of these coefficients suggest that the WASI IQ scores are relatively free of measurement error (The Psychological Corporation, 1999).

The test-retest method was used to determine the stability of WASI scores over a test-retest interval ranging from 2 to 12 weeks ($M = 31$ days). The average coefficients for the child sample were .92 for VIQ, .88 for PIQ, .93 for FSIQ-4, and .85 for FSIQ-2. Due to practice effects, the scores on the second administration were consistently higher than those of the first administration. For the child sample, the IQ scores increased about 2.6 to 5.8 points. The correlation coefficients for the adult sample were similar, and the averages were as follows: .92 for VIQ, .87 for PIQ, .92 for FSIQ-4, and .88 for FSIQ-2. Again, practice effects were observed, and average increases for the adult sample ranged from 1.8 to 3.9 points. All of these coefficients were high for this type of reliability, which indicated that WASI scores have sufficient stability across a short time interval for all age ranges (The Psychological Corporation, 1999).

A third type of reliability assessed during the WASI standardization was interscorer agreement. Agreement for Block Design and Matrix Reasoning are naturally high due to their objective scoring. However, some subjectivity is inherent when scoring Vocabulary and Similarities. The interrater reliability coefficients for Vocabulary and Similarities were .98 and .99, respectively. Although some judgment is needed to score these subtests, these coefficients suggest that they can be scored reliably (The Psychological Corporation, 1999).

Three comparison studies with the WASI and the WISC-III, WAIS-III, and the Wechsler Individual Achievement Test (WIAT; The Psychological Corporation, 1992) were conducted to establish convergent and divergent validity. The correlation coefficient between the WASI FSIQ-4 and WISC-III FSIQ was .87, and the correlation between the WASI FSIQ-2 and WISC-III FSIQ was .81. The correlation coefficient between the two tests regarding VIQ and PIQ was .82 and .76, respectively. With respect to the WAIS-III, the correlation coefficient between the WASI FSIQ-4 and WAIS-III

FSIQ was .92, and the correlation for the WASI FSIQ-2 was .87. For the VIQ and PIQ, the correlation coefficients between the WASI and WISC-III were .88 and .84, respectively. The strength of these correlation coefficients suggests that the IQ scales of the WASI are measuring constructs similar to their WICS-III and WAIS-III counterparts (The Psychological Corporation, 1999).

Axelrod (2002) assessed the validity of the WASI in estimating VIQ, PIQ, and FSIQ scores of the WAIS-III in a group of 72 older males who were seen for neuropsychological evaluations. Correlations of WAIS-III scores with WASI scores were consistently lower than those reported in the WASI manual. In addition, correlations between the WAIS-III and various WAIS-III short forms were higher than that of the WASI and WAIS-III. The correlation coefficient between the WASI FSIQ-4 and WAIS-III FSIQ was .82, and the correlation for the WASI FSIQ-2 was .71. For VIQ and PIQ, the correlation coefficient between the WASI and WAIS-III was .74 and .75, respectively. Furthermore, only 30% of WASI summary scores fell within one standard error of measurement of their WAIS-III counterparts. Axelrod (2002) noted that the lower correlations might be attributed to a relatively small sample size ($N = 72$) or that 37 patients had been diagnosed with a neurological disorder. Dewitt (2003) found correlations between the WASI and WAIS-III to be similar to those presented in the WASI manual. However, like other studies (e.g. Axelrod, 2002) Dewitt found that various WAIS-III short forms provided better estimates of WAIS-III scores. However, it has been noted that correlations between short forms and more comprehensive tests tend to be spuriously high (Silverstein, 1990).

Hays, Reas, and Shaw (2002) examined the concurrent validity of the WASI and K-BIT among psychiatric inpatients. The correlation between the WASI FSIQ-4 and the K-BIT Composite IQ was .89. The K-BIT Vocabulary subtest correlated highest with the WASI Vocabulary subtest ($r = .83$), and the K-BIT Matrices subtest correlated highest with the WASI Matrix reasoning subtest ($r = .83$). The amount of common variance between these tests suggests that they are measuring similar constructs, which provides evidence for the convergent validity of the WASI with another brief measure of intelligence.

Academic achievement and intelligence are related but not synonymous. Therefore, one would expect the correlation coefficients between an intelligence test and an achievement test to be moderate. When the WASI FSIQ-4 was compared to the WIAT Reading and Writing Composites, the resulting correlation coefficient was .72. When The WASI FSIQ-4 was compared to the WIAT Math and Language Composites the correlation coefficient was .64. These correlations are similar to those between the WISC-III and WIAT and the WAIS-III and WIAT. As was expected, the moderate strength of the association between the WASI and WIAT suggests that the two instruments are measuring related but different constructs (The Psychological Corporation, 1999).

The construct validity of the WASI was partially established through the examination of the intercorrelations of the WASI subtests. Statistically, all subtest intercorrelations were significant, which supports the notion of a general intelligence factor. Moreover, the intercorrelations provided evidence of convergence and divergence among the subtests. The overall correlation between Vocabulary and Similarities was

.75. The correlations between Vocabulary and Block Design; and Vocabulary and Matrix Reasoning were .50 and .56, respectively. Block Design correlated highest with Matrix reasoning (r = .59). A correlation of .51 was found between Block Design and Similarities. The correlation between Similarities and Matrix Reasoning was .54. These intercorrelations support the belief that subtests measuring similar constructs correlate more highly with each other than with tests measuring different types of functioning (The Psychological Corporation, 1999).

Exploratory and confirmatory factor analyses were conducted to determine if the WASI subtests measured the constructs of verbal and nonverbal intelligence. The exploratory factor analysis was a joint analysis involving the previously mentioned WISC-III/WASI and WAIS-III/WASI comparison studies. The results supported a factor pattern dividing the verbal and nonverbal subtests, which provided further evidence for the construct validity of the WASI. Once the factor pattern was specified by the joint exploratory analyses, confirmatory analyses were performed with the data from the WASI standardization sample. All analyses compared a general, one-factor model with a Verbal/Nonverbal, two factor model. The goodness-of-fit analyses supported the two-factor model for all age ranges. The root mean squared residual index (RMSR) was 0.517 (The Psychological Corporation, 1999).

Wide Range Intelligence Test (WRIT)

Another brief intelligence test is the Wide Range Intelligence Test (WRIT; Glutting, Adams, & Sheslow, 2000). Like the K-BIT and WASI, the WRIT is an efficient, psychometrically sound instrument. It is a 30 minute, individually administered test for persons ranging from 4 to 85 years. Norms from the WRIT are representative of the United States population and linked to the Wide Range Achievement Test 3 (WRAT 3; Wilkinson, 1993). The WRIT uses a three-level model to describe the organization of abilities. At the first level is general intelligence. At the next level are the constructs of fluid and crystallized intelligence, like the K-BIT and WASI. The third level of the WRIT consists of four subtests (Matrices, Verbal Analogies, Diamonds, and Vocabulary) that have been shown to be strongly associated with general intelligence.

The Matrices subtest uses a traditional matrix analogy format like that of the WAIS-III and K-BIT. This subtest is a measure of fluid intelligence. The other measure of fluid intelligence is the Diamonds subtest, which is a timed construction task requiring spatial skills. It was adapted from the Kohs Block Design Test (Kohs, 1927). Diamonds involves the use of diamond-shaped puzzle pieces (2 common-border diamonds shaped as a "V", or 3 common-border diamonds shaped as a "lightening bolt"). Participants then use these pieces to construct a picture. Verbal Analogies is a measure of crystallized intelligence and requires verbal abstraction and generalization of meaning from the examinee. In this subtest individuals say a word that best completes an analogy. Vocabulary completes the WRIT's Verbal IQ scale. Vocabulary tasks are present in most IQ tests (Glutting, Adams, & Sheslow, 2000).

As mentioned previously, the WRIT is a psychometrically sound instrument. The technical data in the WRIT manual shows that the standardization sample consisted of 2,285 individuals ranging from 4 to 85 years of age. Coefficient alpha was used to determine internal-consistency reliability. The mean coefficients for the WRIT's General, Verbal, and Visual IQs were .95, .94, and .92, respectively. In addition, person separation reliabilities were established. Like other measures of internal consistency, person separation reliabilities estimate measurement error. In the overall sample, person separations ranged from .94 to .97 for the four subtests. The reliabilities obtained by coefficient alpha and person separations indicated that the WRIT was relatively free of measurement error. Item separation reliabilities were also determined, which indicated how well the items defined the variables being measured. For the entire sample, the item separation values were 1.00 for all four subtests. Such high values established that subtest items were sufficiently separated from easy to hard (Glutting, Adams, & Sheslow, 2000).

The test-retest method was used to determine the stability of WRIT scores over a short timeframe. The test-retest interval ranged from 6 to 115 days ($M = 30.5$ days). The average coefficients for the total sample were .96 for Verbal IQ, .90 for Visual IQ, and .96 for General IQ. Not unlike other IQ tests, the scores on the second administration were higher than those of the first administration. For the total sample, the General IQ scale increased about 5.7 points on average, whereas the Verbal and Visual scales increased 4.5 points and 6.6 points, respectively. These coefficients indicated that WRIT scores had sufficient stability across a short time interval for all age ranges (Glutting, Adams, & Sheslow, 2000).

Interscorer agreement was also assessed. When scoring Matrices and Diamonds little judgment was needed. However, Vocabulary and Verbal Analogies require judgment by the examiner. Intraclass correlation coefficients were calculated to determine interscorer reliability. This coefficient determined whether an association was present among raters, as well as if quantitative values on ratings were the same. For the Vocabulary and Verbal Analogies subtests, the intraclass correlation coefficients were .98

and .99, respectively. Consequently, very little error was associated with scoring these subtests (Glutting, Adams, & Sheslow, 2000).

The validity of the WRIT was examined both internally and externally. To determine internal validity, factor analyses were conducted. An exploratory analysis using the entire standardization sample was utilized via principal components method. Results indicated that a one-factor (general intelligence) model was the only one that satisfied all criteria for factor retention. Once the correct number of factors was established and because retention of a single factor rules out rotation, an unrotated principal components analysis was conducted. The results indicated that all subtests loaded highly with the single factor, which means that WRIT subtest scores were best conceptualized as measuring the construct of general intelligence (Glutting, Adams, & Sheslow, 2000).

Next, a forced, two-factor solution (principal axis factor analysis with promax rotation) was attempted to determine if the subtests would load with their hypothesized constructs. The two factors resulted in the hypothesized fluid/crystallized dichotomy. Furthermore, the two-factor solution accounted for 19% more variance among WRIT subtests. This suggests that the WRIT provides a secondary contribution other than general intelligence. However, the relationship between the Verbal factor and Visual factor was also assessed. The two factors had a correlation coefficient of .75 within the forced, two-factor solution. These results are similar to other IQ measures like WAIS-III and WISC-IV. Consequently, this strong relationship supports the earlier hypothesis regarding a single, general intelligence factor. On the other hand, the one- and two-factor solutions were further studied through confirmatory factor analysis, and results of the goodness-of-fit analyses showed a preference for the two-factor solution (Glutting, Adams, & Sheslow, 2000).

Much like the WASI, joint factor analyses were also conducted with the WRIT/WISC-III and WRIT/WAIS-III. These joint analyses also supported a two-factor solution. Furthermore, they showed that since there was so much redundancy in what the WRIT was measuring relative to the WISC-III and WAIS-III, constructs between the WRIT and the two Wechsler scales could serve as substitutes for each other. In general, the results of all these factor analyses support the construct validity of the WRIT's Verbal, Visual, and General Intelligence Scales (Glutting, Adams, & Sheslow, 2000).

Three comparison studies with the WRIT and the WISC-III ($N = 100$), WAIS-III ($N = 100$), and the WRAT 3 ($N = 1182$) were conducted to establish convergent and divergent validity, measures of external validity. The correlation coefficient between the WRIT General IQ and WISC-III FSIQ was .90. The correlation coefficient between the WISC-III VIQ and WRIT Verbal IQ was .85, and the WISC-III PIQ and WRIT Visual IQ correlation was .78. With respect to the WAIS-III, the correlation coefficient between the WRIT and WAIS-III FSIQ was .91. The correlation coefficient between the WAIS-III VIQ and WRIT Verbal IQ was .90, and the WAIS-III PIQ and WRIT Visual IQ correlation was .85. The strength of these correlation coefficients suggests that the IQ scales of the WRIT are measuring constructs similar to the WICS-III and WAIS-III (Glutting, Adams, & Sheslow, 2000).

Again, academic achievement and intelligence are related but not synonymous. When the WRIT General IQ Scale was compared to WRAT 3 Reading subtest standard scores, the resulting correlations coefficients ranged from .45 to .63 across the age bands.

When the WRIT General IQ was compared to WRAT 3 Math standard scores, correlations ranged from .46 to .58 and from .46 to .54 for the Spelling subtest standard scores. As expected, the moderate strength of the association between the WRIT General IQ Scale and WRAT 3 subtest standard scores suggests that the two instruments are measuring related but different constructs (Glutting, Adams, & Sheslow, 2000).

Shields, Konold, and Glutting (2004) compared the WRIT's utility in predicting WRAT 3 performance. In that study the WRIT's differential validity was assessed across race/ethnicity, gender, and education level. Of the 90 simultaneous tests, 65 revealed no statistically significant between-group differences. Furthermore, the majority of statistically significant differences were found to have little practical influence when measures of effect size were taken into consideration. This supports the contention that the WRIT does not possess bias in the prediction of academic achievement.

Purpose of the Study

There are currently no published comparisons of the WASI and WRIT. In addition to the studies presented in the test manuals, no published validity studies have been conducted for the WRIT, and few validity studies (e.g. Axelrod, 2002; Hays et al., 2002; and Dewitt, 2003) have been conducted for the WASI. With the exception of Hays et al., all of the validity studies regarding the WRIT and WASI have examined the relationship of these brief tests with comprehensive measures of intelligence. Since there is only one validation study for either of these instruments with other brief measures of intelligence, one purpose of this study was to further verify the relationship between the WASI and the WRIT by replicating Collins' (2002) study regarding their convergent validity.

Collins (2002) examined the convergent validity of the WRIT and the WASI among a small sample ($N = 66$) of elementary, middle, and high school students. The correlations were similar to those between the WRIT and the more comprehensive tests previously mentioned. The correlation coefficients between WASI and WRIT IQ scores and subtest scores can be viewed in Table 2. The correlations between the WRIT General IQ and WASI FSIQ-2 and FSIQ-4 were .83 and .85, respectively. The Verbal IQ scale of the WRIT correlated highly with the VIQ of the WASI ($r = .82$), and the Visual IQ scale of the WRIT correlated highly with the WASI PIQ ($r = .78$).

Collins also examined the correlations between the corresponding subtests. The highest correlation was between the Vocabulary subtests ($r = .80$). The correlation between the WRIT Diamonds subtest and the WASI Block Design subtest of was .71. The correlation between the WRIT Matrices subtest and the WASI Matrix Reasoning subtest was .69, and the correlation between WRIT Verbal Analogies subtest and the WASI Similarities subtest was .68. It was concluded that both tests were measuring similar constructs, and that both were good indicators of general intelligence.

By replicating the Collins' (2002) study the relationship between the WASI and WRIT can be further verified. Furthermore, by adding to the sample size of that study ($N = 66$), a more accurate representation of the relationship between the WASI and WRIT may be provided through additional analyses such as factor analyses.

There were two hypotheses regarding the first purpose of the study. The WASI and WRIT test manuals both provide evidence supporting a fluid-crystallized dichotomy and the ability to measure the construct of intelligence. Furthermore, Collins (2002) reported that the convergent validity between the like scales of the WASI and WRIT was significant. Therefore, it was hypothesized that the similar IQ scales and subtests of the WRIT and WASI would be highly correlated since there is evidence supporting that they measure the same or similar constructs. Moreover, all subtests and IQ scales of the WRIT and WASI are measuring general intelligence to some extent. However, verbal subtests focus on measuring crystallized intelligence, and nonverbal subtests center on fluid intelligence. Therefore, a second hypothesis is that dissimilar subtests and IQ scales will be moderately correlated.

A second purpose of this study was to examine the factorial validity of these two tests by examining the joint factor structure of the WASI and WRIT. In similar analyses between the WRIT and WISC-III; and WRIT and WAIS-III, joint factor analyses offered strong support that the two instruments measured virtually indistinguishable latent dimensions (Glutting et al., 2000), and measured nearly identical constructs. Consequently, that information provided evidence for the theoretical underpinnings of the WRIT. Based on these results, it was hypothesized that joint exploratory factor analyses of the WASI and WRIT would result in support for a general intelligence factor, as well as separate crystallized/verbal and fluid/nonverbal factors.

Method

Participants

Table 1 presents the sample demographic characteristics for the 86 individuals who participated in the study. The participants ranged from first graders to adults with a Master's degree ($N = 8$) and were either unpaid volunteers, children of parents who had provided consent for participation in the study, or students who had been referred for a psychological evaluation. Referred students were limited to those in need of reevaluations, since initial evaluations require more comprehensive measurement of intelligence. The distribution of gender was predominately female. Due to the ethnic makeup of the geographic area where data were collected, the sample was predominately Caucasian but included a small number Hispanic and Asian participants. Participants ranged in age from 6.17 to 53.75 years ($M = 11.83$, $SD = 7.58$). The sample was also dominated by students without disabilities, but it included two students who had been diagnosed as learning disabled, one learning and emotionally disabled, and one Other Health Impaired.

Instruments

Wide Range Intelligence Test (WRIT)

The WRIT is a brief measure of intelligence designed for children and adults between the ages of 4 and 85 and can be administered in less than 30 minutes. The WRIT consists of a record form, a manual, a stimulus manual, and manipulatives (i.e. Diamond Chips). A stopwatch is also necessary for administration. The WRIT consists of four subtests. The Verbal Analogies subtest is a sentence completion task. The Vocabulary subtest requires individuals to provide definitions. These subtests combine to measure Verbal IQ. The Matrices subtest is a pattern completion task, and the Diamonds subtest requires subjects to recreate a stimulus pattern. These combine to measure Visual IQ. All four subtests combine to measure the WRIT General IQ. Global scores are reported in terms of an intelligence quotient (IQ) with a mean of 100 and a standard deviation of 15. Subtest scores are also reported on a scale with a mean of 100 and standard deviation of 15.

Wechsler Abbreviated Scale of Intelligence (WASI)

The WASI is a brief measure of intelligence for individuals between the ages of 6 and 89. The WASI requires approximately 30 minutes to administer when using the full battery (FSIQ-4) and 15 minutes when using the abbreviated battery (FSIQ-2). The WASI consists of a record form, a manual, a stimulus manual, and manipulatives (i.e. Pattern Blocks). A stopwatch is also necessary for administration. The full battery consists of four subtests. The Vocabulary subtest requires individuals to provide definitions. The Similarities subtest requires an individual to describe a connection between objects or concepts. The combination of these two tests produces the WASI VIQ. The Block design subtest involves recreating a stimulus pattern, while the Matrix Reasoning subtest is a pattern completion task. These two subtests combine to make the WASI PIQ. The combination of all four subtests produces the WASI FSIQ-4. The abbreviated battery (FSIQ-2) is the combination of the scores obtained from only the Vocabulary subtest and Matrix Reasoning subtest. Global scores are reported in terms of an intelligence quotient (IQ) with a mean of 100 and a standard deviation of 15. The subtest scores for the WASI are reported in terms of a *T* score with a mean of 50 and a standard deviation of 10.

Procedure

Permission was obtained from the principals of the schools where the study took place. The teachers were then informed about the study and asked for their assistance identifying children for participation in the study (Appendix A). Teachers sent permission forms (Appendix B) home with the students for their parents to sign and return. Participants were chosen from the group of students whose parent or legal guardian granted permission. Adults who wished to contribute to the study were allowed to participate. Adults who participated in the study were either family or friends of the test administrator. Archival data from students who had been administered the WASI and WRIT in a psychological evaluation were utilized, as well.

During testing, the two tests were administered in random counterbalanced order to control for possible order effects. Each individual was tested during a single session. One examiner administered all the tests. The administrator was a school psychologist intern professionally trained in psychometric testing, who conducted testing in a manner consistent with professional practice.

Data Analysis

In order to determine the convergent validity of the WASI and WRIT, WASI subtest, *T* scores were converted to standard scores ($M = 100$, $SD = 15$) so that the subtest scores on both tests were in the same measurement unit. Pearson product-moment correlation coefficients were calculated to assess the levels of convergent validity between the various subtests and composites of the WASI and WRIT. Dependent *t*-tests for differences between means were also used to examine *level* differences (McDermott, 1988) between similar subtests and composite scores of the WASI and WRIT. To determine if differences were meaningfully different, effect size estimates were calculated using Glass' Δ (Glass & Hopkins, 1996). The test scores of the 86 participants in the present study were used to calculate the correlation coefficients, *t*-scores, and effect size estimates.

Collins (2002) data set was merged with the present data set in order to conduct an exploratory factor analysis. The demographic information for the participants in Collins (2002) can be viewed in Table 1. The samples were merged because, independently, neither had adequate sample size for a factor analysis, but merging the sets produced acceptable sample size for factor analysis given high communality estimates.

The relationship between the latent dimensions of the WASI and WRIT was analyzed through joint exploratory factor analysis with multiple criteria for factor extraction. These criteria included eigenvalues greater than one, the scree test, parallel analysis, and examination of theoretical convergence.

Principal axis factor analysis with promax (oblique) rotation when multiple factors were extracted was conducted using SPSS 11.0 for Mac. Parallel analysis is based on a comparison of eigenvalues obtained from sample data to eigenvalues obtained from random data. A model is specified with the same number of common factors as real eigenvalues that are greater than the eigenvalues expected from random data (Fabrigar, Wegener, MacCallum, and Strahan, 1999). Parallel analysis was conducted using the *Monte Carlo PCA for Parallel Analysis* computer program (Watkins, 2000) with 100 replications to provide stable eigenvalue estimates.

Results

Global Scale Comparisons

Table 2 presents the correlations and r^2s between the IQ scores of the WASI and WRIT. The test scores of the 86 participants in the present study were used to calculate the correlations. All correlations between the IQ scores were statistically significant ($p <$

.001). The correlations between the corresponding IQ scores between the WASI and WRIT ranged from .76 to .82. The correlation between the WASI FSIQ-4 and the WRIT General IQ was .82. This correlation was the strongest of all the comparisons and accounted for a large amount of shared variance ($r^2 = .67$). The Correlation between the Verbal IQs of the WASI and WRIT was .77, while the correlation between the WASI PIQ and the WRIT Visual IQ was .76.

The correlations regarding the fluid / crystallized dichotomy within and between the two measures were low to moderate. The correlation between the WASI VIQ and WASI PIQ was .54. The correlation between the WASI VIQ and WRIT Visual IQ was similar ($r = .51$) and accounted for only 26% shared variance. The correlation between the WRIT Verbal IQ and WRIT Visual IQ was .49, while the correlation between the WRIT Verbal IQ and WASI PIQ was .43 and accounted for a smallest amount of shared variance ($r^2 = .18$) among all IQ comparisons.

Table 3 presents the dependent *t*-test results for similar IQ scales of the WASI and WRIT. Participants obtained statistically equivalent WASI FSIQ-4 and WRIT General IQ scores $t(85) = -1.01, p = .315$. The difference between the means of WASI VIQ and WRIT Verbal IQ was not significant $t(85) = 0.47, p = .637$, as was with the WASI PIQ and WRIT Visual IQ $t(85) = -1.76, p = .081$.

Subtest Comparisons

Table 2 presents the correlations between the various subtests of the WASI and WRIT. The test scores of the 86 participants in the present study were used to calculate the correlations. All correlations between the corresponding subtests were statistically significant ($p < .001$) and ranged from .55 to .73. The strongest correlation among the subtests was for the Vocabulary scores of the WASI and WRIT ($r = .73$), which accounted for 53% shared variance. The other correlations between the corresponding subtests of the WASI and WRIT were moderately correlated. The correlation between the WASI Matrix Reasoning subtest and the WRIT Matrices subtest was .67. The correlation between the WASI Similarities subtest and WRIT Verbal Analogies subtest was .58. The lowest correlation between corresponding subtests was .55 for the WASI Block Design and WRIT Diamonds subtest comparison.

The correlations between the dissimilar subtests of the WASI and WRIT were the lowest among all subtest comparisons. Dissimilar subtest comparisons included those between verbal and non-verbal subtests. Although moderate to low in strength, all comparisons between the dissimilar subtests were statistically significant ($p < .05$), and the correlations ranged from .26 to .48. The strongest correlation was between the WASI Vocabulary subtest and the WRIT Matrices subtest ($r = .48$). The correlation between The WASI Similarities subtest and WRIT Matrices subtest was also .48. The weakest correlation among dissimilar subtests was between the WASI Vocabulary subtest and the WRIT Diamonds subtest ($r = .26$). The relationship between these two subtests accounted for only 7% shared variance.

Table 3 presents the dependent *t*-test results for comparisons between similar subtests of the WASI and WRIT. Students obtained statistically equivalent WASI Block Design and WRIT Diamonds subtest scores $t(85) = -0.95, p = .343$. The difference between the mean subtest scores of the WASI Matrix Reasoning and WRIT Matrices subtest also was not significant $t(85) = -0.52, p = .603$.

Differences between means of the remaining corresponding subtests were significant. On average, participants scored significantly higher on the WRIT Vocabulary subtest (M = 107.83, SD = 11.68) than on the WASI Vocabulary subtest (M = 104.76, SD = 13.68), $t(85) = -3.10$, $p = .003$, $\Delta = .21$. Participants also scored significantly higher on the WASI Similarities subtest (M = 110.80, SD = 14.01) relative to their performance on the WRIT Verbal Analogies subtest (M = 106.77, SD = 12.04), $t(85) = 3.09$, $p = .003$, $\Delta = .27$. However, these mean differences were well within the standard error of measurement of both measures and also represented small effect sizes (Cohen, 1992) based on Glass' Δ (Glass & Hopkins, 1996). These small effect sizes suggested that the significant differences between the means were not clinically significant.

Joint Exploratory Factor Analysis

A joint exploratory factor analysis could not be conducted on the data obtained from the sample in the present study. The communality estimates were not at a high enough magnitude considering the small sample size (N = 86). Therefore, the data obtained from the sample in the current study were combined with the data obtained from the sample by Collins (2002) to obtain an adequate sample size (N = 152) for an exploratory factor analysis.

Pearson product-moment correlations, promax structure coefficients, eigenvalues, and the percent of variance accounted for are presented in Table 4. The Kaiser-Meyer-Olkin Measure of Sampling Adequacy was .90 and Bartlett's Test of Sphericity was 887.52, $p < .0001$. Communality estimates ranged from .55 to .77 (Mdn = .63). The first exploratory analysis employed principal axis methodology. Only one factor had an eigenvalue greater than one (see Table 4). Based on eigenvalues >1 and parallel analysis, (see figure 1), only one factor (general ability "g") was extracted through the initial principal axis factor analysis.

The results of the scree test and examination of theoretical convergence supported a two factor solution (see Figure 1). Based on these results, a forced, two-factor solution was conducted. This principal axis factor analysis with promax rotation was performed to determine whether the WASI Vocabulary, WASI Similarities, WRIT Vocabulary, and WRIT Verbal Analogies subtests would correspond to one factor and the WASI Block Design, WASI Matrix Reasoning, WRIT Diamonds, and WRIT Matrices on the second factor. Table 4 presents the factor structure coefficients for each subtest. Results indicated that the WASI Vocabulary, WASI Similarities, WRIT Vocabulary, and WRIT Verbal Analogies had a higher factor structure coefficients on the first factor; whereas the WASI Block Design, WASI Matrix Reasoning, WRIT Diamonds, and WRIT Matrices had higher factor structure coefficients on the second factor.

The subtests that were associated with the first factor displayed a unified pattern of verbal and crystallized cognitive skills. Conversely, the subtests that were associated with the second factor portrayed a high degree of visual and fluid abilities. Consequently, the two factors captured more variation among WASI/WRIT performance than the one-factor solution. Specifically, the second factor accounted for 10.50% more variance than the one-factor model (77.08% for the two-factor solution vs. 66.62% for the one-factor solution).

Although the two-factor model matched the theoretical expectations for subtest associations with factors, there were cross-loading problems in the two-factor model (see

Table 4). Subtests had higher factor structure coefficients with their theoretically consistent dimension, but all subtests also had salient (≥ .40) factor structure coefficients on the alternate factor illustrating cross-loading problems. The strength of the relationship between the two factors was assessed. Factor I (crystallized/verbal) and Factor II (fluid/nonverbal) correlated at .75. The magnitude of this coefficient indicated that scores from the two factors share 56% of their variance.

Discussion

Global Scale Results

The present study examined the construct validity of two newer brief intelligence tests, the Wechsler Abbreviated Scale of Intelligence (WASI) and the Wide Range Intelligence Test (WRIT) with a sample of individuals ranging from elementary students to adults. As hypothesized, positive correlations ranging from moderate to high were found between the various similar scales of the WASI and WRIT, and convergent validity among the corresponding scales of the WASI and WRIT was further established through the replication of results obtained by Collins (2002).

The correlation between the WASI FSIQ-4 and the WRIT General IQ was statistically significant and high with 67% shared variance, indicating the measurement of the same construct (general intelligence). The correlations between the WASI VIQ and WRIT Verbal IQ; and WASI PIQ and WRIT Visual IQ were also strong, 59% and 58% shared variance, respectively.

Table 2 shows the correlation coefficients between the corresponding IQ scores of the WASI and WRIT in the present study and those in Collins (2002). The correlations between the corresponding IQ scores of the WASI and WRIT in this study were similar to those found by Collins (2002). The correlations between the corresponding IQ scores of the WASI and WRIT in this study were also similar to the correlations each test had with comprehensive IQ tests as reported in the WASI and WRIT manuals (WASI; Psychological Corporation, 1999; WRIT; Glutting et al., 2000), and those found by Canivez (1995) between the K-BIT and WISC-III. When compared to previous research regarding the construct validity of intelligence tests, the results of the current study support the concurrent validity of the WASI and WRIT as brief measures of intelligence.

Since they purport to measure different constructs, the weaker correlations of the fluid/crystallized dichotomy between the two measures was expected. The correlation between WASI VIQ and WRIT Visual IQ ($r = .51$) was lower than the convergent associations previously mentioned, as was the correlation between the WASI PIQ and WRIT Verbal IQ ($r = .43$). These correlations were lower than those reported by Collins (2002), where the correlations between the WASI VIQ and WRIT Visual IQ; and WASI PIQ and WRIT Verbal IQ were .67 and .65, respectively.

However, the correlation coefficients in the present study were still moderately high, despite the fact that verbal and non-verbal IQ scales purport to measure different constructs. Due to their strong associations with the general intelligence factor or "g", the moderate correlations between the dissimilar IQ scales of the WASI and WRIT may be explained by Macmann and Barnett's (1994) contention that verbal and performance factors could be described as "truncated or degraded versions of the general factor" (Macmann & Barnett, 1994, p. 180).

Mean scores of corresponding IQ scales between the WASI and WRIT were compared in order to assess if participants scored significantly better on one scale than the other. The results indicated that participants scored statistically equivalent on all corresponding IQ scales of the WASI and WRIT. These results were different than those reported by Collins (2002). Collins reported a statistically significant difference between the mean scores of the WASI PIQ and WRIT Visual IQ, $t(65) = 2.50, p = .015$. However, the difference Collins (2002) reported represented a small effect size (Cohen, 1992) based on Glass' Δ (Glass & Hopkins, 1996), $\Delta = .24$, which suggested that the difference between the means was not clinically significant.

Furthermore, the mean differences between the IQ scales of the WASI and WISC-III; and the WRIT and the WISC-III, as reported in their respective manuals (WASI; Psychological Corporation, 1999; WRIT; Glutting et al., 2000), were small and similar to those found between the WASI and WRIT in the current study. The results of the current study indicated that, in general, participants scored equivalently on the corresponding IQ scales of the WASI and WRIT, which further supports the concurrent validity of the WASI and WRIT as brief measures of intelligence.

Subtest Comparison Results

At the subtest level, all WASI subtests were significantly correlated with all subtests of the WRIT. The correlations among the corresponding subtests ranged from .55 to .73 and were lower than those between the corresponding IQ scales. The highest correlation was between the Vocabulary subtests, where there was 53% shared variance. This suggested that the two subtests were measuring similar constructs. The WASI Matrix Reasoning and WRIT Matrices subtest experienced a moderately strong correlation with 45% shared variance, which suggested that they were measuring similar constructs, as well.

The other corresponding subtests had weaker correlations. The WASI Similarities subtest and the WRIT Verbal Analogies subtest had a correlation of .58, which resulted in a moderate amount of shared variance ($r^2 = .34$). This weaker correlation may be explained by the fact that the subtests are different tasks. The Similarities subtest requires an individual to link two presented items conceptually, where the Verbal Analogies subtest presents an individual with a relationship and requires that individual to create and complete the same relationship with different information.

The WASI Block Design subtest and the WRIT Diamonds subtest had a correlation of .55, which also resulted in a moderate amount of shared variance ($r^2 = .30$). This weaker correlation may be explained by the increased amount of abstraction inherent in the Diamonds subtest. Both subtests are perceptual/organizational tasks that require an individual to recreate a visually presented pattern with manipulatives. However, the Block Design subtest is, perceptually, always two-dimensional and in the form of either a two-by-two or three-by-three grid. The Diamonds subtest may be more abstract. The Diamonds subtest becomes, perceptually, three-dimensional and never remains in a set form.

The correlations between the corresponding subtests of the WASI and WRIT were similar to some of those reported by Collins (2002). Collins (2002) highest corresponding subtest correlation was also between the WASI and WRIT Vocabulary subtests, and the correlation between the WASI Matrix Reasoning and the WRIT Matrices subtest was consistent with the results of the present study, as well (see Table 2). However, the correlations reported by Collins (2002) between the WASI Similarities and WRIT Verbal Analogies subtest; and WASI Block Design and WRIT Diamonds subtest were stronger (see Table 2). In addition, the corresponding subtest correlations between the WASI and WRIT in the present study were similar to those found between the WASI and WISC-III; and WASI and WAIS-III (The Psychological Corporation, 1999).

The correlations among the dissimilar subtests of the WASI and WRIT were lower than those for corresponding subtests, as would be expected given the different constructs being assessed by the various subtests. As was previously mentioned, dissimilar subtest comparisons included those between verbal and non-verbal subtests. All the correlations among the dissimilar subtests were significant ($p < .05$) and ranged from .26 to .48. The strongest correlations were between the WASI Vocabulary subtest and the WRIT Matrices subtest ($r = .48$) and the WASI Similarities subtest and WRIT Matrices subtest ($r = .48$). The weakest correlation among dissimilar subtests was between the WASI Vocabulary subtest and the WRIT Diamonds subtest ($r = .26$). The relationship between these two subtests accounted for only 7% shared variance. These weaker correlations suggested that the WASI and WRIT might be measuring more than one construct. However, all dissimilar subtest correlations were still positive and most were moderate in strength, which suggested that a unifying construct "g" was being measured in all subtests.

The results of the current study regarding dissimilar subtests were slightly different than those reported by Collins (2002). In fact, the weakest correlation between dissimilar subtests reported by Collins (2002) was still stronger than the strongest correlation between dissimilar subtests in the current study (see Table 2). Collins (2002) reported dissimilar subtest correlations between the WASI and WRIT that ranged from .50 to .62.

Mean scores of corresponding subtests between the WASI and WRIT were compared in order to assess if participants scored significantly better on one subtest than the other. The results indicated that participants scored statistically equivalent on WASI Block Design and WRIT Diamonds subtest; and WASI Matrix Reasoning and WRIT Matrices subtest. However, subjects scored significantly higher on the WRIT Vocabulary than on the WASI Vocabulary subtest ($\Delta = .21$). Subjects also scored significantly higher on the WASI Similarities subtest relative to their performance on the WRIT Verbal Analogies subtest ($\Delta = .27$). Nonetheless, these mean differences represented small effect sizes (Cohen, 1992) based on Glass' Δ (Glass & Hopkins, 1996), which suggested that the differences between the means were not clinically significant or of practical concern.

These results were similar to those reported by Collins (2002). Although Collins (2002) reported that subjects scored significantly higher on the WRIT Diamonds subtest relative to their performance on the WASI Block Design subtest, the mean difference represented a small effect size ($\Delta = .33$), thus, not clinically significant. Unfortunately, the WASI and WRIT manuals do not provide comparison statistics for mean differences.

Exploratory Joint Factor Analysis

There was a concern that the sample size for the joint factor analysis would be too small. However, when communalities are consistently high ($\geq .60$) the detrimental effects of sampling receive a low weight (MacCallum, Widaman, Zhang, & Hong, 1999). The commonalities in the current study were as follows: WASI Vocabulary = .77, WASI

Block Design = .56, WASI Similarities = .69, WASI Matrix Reasoning = .63, WRIT Verbal Analogies = .63, WRIT Vocabulary = .73, WRIT Matrices = .64, and WRIT Diamonds = .55.

The majority of the communalities in the current joint factor analysis were high, but two were in the range of .50. Nonetheless, when communalities are in the range of .50, it is still not hard to get good recovery of population factors, but one must have well-determined factors and a somewhat larger sample in the range of 100 to 200 (MacCallum et al., 1999). The current study had four variables representing each common factor, which supports the research regarding "overdetermination" of common factors (MacCallum, et al., 1999). Moreover, the sample size in the present joint factor analysis ($N = 152$) was in the range of 100 to 200. These results indicated that the sample in the current study provided for good recovery of population factors.

The results of the first exploratory factor analysis in the current study supported the extraction of only one factor (general intelligence "g"). These results were different than the results reported in the WASI manual regarding joint factor analyses with the WASI/WISC-III ($N = 176$) and WASI/WAIS-III ($N = 248$). In these exploratory, joint-factor analyses (principal axis methodology with Promax rotation) the WASI Vocabulary and WASI Similarities subtests loaded high on one factor (.90 and .66, respectively) and extremely low on all other factors. Likewise, the WASI Block Design and WASI Matrix Reasoning subtests loaded high on a second factor (.73 and .68, respectively) and very low on all other factors (WASI: The Psychological Corporation, 1999).

Consequently, the results reported in the WASI manual supported the premise that the two-factor model best fit the data for the total sample (The Psychological

Corporation, 1999). The differences just discussed could be attributed to the more diverse, representative sample of the analyses reported in the manual, or the fact that those analyses included additional subtests of the WISC-III and WASI-III. It is possible that the nature of the subtests included created variance in a way that made it easier to get good recovery of population factors (MacCallum et al., 1999).

The results of the current joint factor analysis were similar to those of the exploratory analysis reported in the WRIT manual. The first exploratory analysis (N = 2,285) employed principal components methodology. Results showed that a one-factor, "g"-based model was the only solution to satisfy all criteria for factor retention (Glutting et al., 2000).

This WRIT's structure, as reported in the manual, was also similar to the structure found with the WASI/WRIT. As reported in the WRIT manual, the WRIT Vocabulary subtest's loading on the general factor was .80, as compared to .83 in the current study. Verbal Analogy's loading was .82, as compared to .79. Diamonds' loading was .64, as compared to .72, and Matrices' loading was .72, as compared to .78. The WASI/WRIT joint exploratory factor analysis and the WRIT exploratory factor analysis reported in the WRIT manual indicated that subtest scores from the WASI and WRIT are best conceptualized as providing strong measurements of the robust construct of general intelligence.

Glutting et al. (2000) argued that the reductive aspect of factor analysis makes it difficult to identify more than one factor with the WRIT. It only consists of four subtests, and a one-factor model has the potential of containing three, or more, indicators per factor. However, the WASI/WRIT analysis provided four indicators for each theorized factor, and the results still supported retention of a single factor ("g"). The results of current study and those reported in the WRIT manual do not support the earlier hypothesis that the WASI and WRIT are measuring intelligence across multiple dimensions. Instead, the results suggest that the two tests are principally measuring a single, higher-order factor ("g").

The forced-two factor solution in the current study was also similar to the forced-two factor solution reported in the WRIT manual. In both analyses all the subtests had a higher loading magnitude on their hypothesized factor. In other words, verbal subtests loaded highest with the verbal/crystallized intelligence factor (Factor I), and the non-verbal subtests loaded highest with the performance/fluid intelligence factor (Factor II). Likewise, the forced, two-factor solution accounted for more subtest-score variance in both analyses. In the current study, the forced, two-factor solution accounted for 10.50% more variation among WASI/WRIT subtest scores. The WRIT manual reported 19% more variation explained (Glutting et al., 2000). These results suggested that there may be some utility in interpreting the WASI and WRIT in terms of a crystallized/fluid dichotomy.

However, both analyses indicated a strong relationship between the two factors. The WRIT manual reported that the correlation between the two dimensions was .75 (Glutting et al., 2000). The same correlation was found in the forced two-factor analysis conducted in the current study. The magnitude of this coefficient indicated that scores from the two factors were redundant and shared 56% of their variance. These results support the results of the initial joint exploratory factor analysis in the current study, which suggested that the WASI and WRIT were best measuring "g." Still, a hierarchical model allows for both to be considered.

Although the results of the current study and the results of the analyses reported in the WRIT manual were very similar, there were differences regarding the cross-loading of structure coefficients in the forced, two-factor model. The results reported in the WRIT manual did not indicate problems with cross-loading. However, in the current study, all structure coefficients cross-loaded (see Table 4). Although this difference could be due to a restricted range in the WASI/WRIT sample, the cross-loading problems further support a conclusion that the WASI and WRIT are best conceptualized as providing strong measurements of general intelligence in the present study. A confirmatory factor analysis should be conducted in the future to see which model is the best fit for these data.

Limitations

There were several limitations to this study that should be addressed in future research. In general, the limitations revolved around the demographics of the sample. The sample size was relatively small ($N = 86$), which could have resulted in greater sampling error. All individuals who participated in the study were from rural, central Illinois and primarily Caucasian. The sample's limited geographic and ethnic diversity restricts the generalization of results to groups with different characteristics. Furthermore, the WASI and WRIT were designed to assess individuals of different ages ranging from very young children to elderly adults. Although a few adults were included in the sample, the vast majority of subjects were between the ages of six and 12 years-old. Utilizing a sample with more adults would provide insight to how these measures compare with a wider range of individuals. Future studies involving these cognitive ability measures should incorporate a more global representation of the population in order to avoid the aforementioned limitations. Moreover, a confirmatory factor analysis should be conducted in the future to determine which model is the best fit for the data.

Conclusions

The results of the current study indicated that the corresponding IQ scales and subtests of the WASI and WRIT were highly correlated. These results were generally consistent with those reported by Collins (2002). Furthermore, the WASI and WRIT manuals (WASI; Psychological Corporation, 1999; WRIT; Glutting et al., 2000) reported high correlations with comprehensive intelligence tests (e.g. WISC-III and WAIS-III) in analyses similar to those used in the current study. These results provided convergent evidence for the construct validity of the WASI and WRIT and supported the contention made by Collins (2002) that psychologists using the WASI and WRIT should be confident that these tests are measuring similar constructs.

However, at the present time, caution should be exercised when interpreting the WASI and WRIT in terms of a fluid/crystallized dichotomy. The results of the joint exploratory factor analysis supported the retention of only one factor (general intelligence "g"). This conclusion was also reported in the WRIT manual (Glutting et al., 2000). A forced, two-factor solution resulted in 10.50% more variation explained among WASI/WRIT subtest scores. However, the two factors correlated at a high magnitude (r = .75), and all subtests cross-loaded with their theoretical dissimilar factor. The results of the factor analysis supported a conclusion that the WASI and WRIT data in this study are best conceptualized as providing strong measurements of general intelligence. However, a confirmatory factor analysis should be conducted in the future to see which model is the best fit for the data.

References

Axelrod, B. N. (2002). Validity of the Wechsler Abbreviated Scale of Intelligence and other very short forms of estimating intellectual functioning. *Assessment, 9*, 17-23.

Canivez, G. L. (1995). Validity of the Kaufman Brief Intelligence Test: Comparisons with the Wechsler Intelligence Scale for Children-Third Edition. *Assessment, 2*, 101-110.

Cohen, J. (1959). The factorial structure of the WISC at ages 7-6, 10-6, and 13-6. *Journal of Consulting and Clinical Psychology, 23*, 285-299.

Cohen, J. (1992). A power primer. *Psychological Bulletin, 112*, 155-159.

Collins, J. (2002). Correlations between the Wide Range Intelligence Test (WRIT) and the Wechsler Abbreviated Scale of Intelligence (WASI): Global and subtest comparisons. thesis (Eastern Illinois University). unpublished, 1-39.

Dewitt, A. (2003). Comparability of short form approaches of the Wechsler Adult Intelligence Scale-III to the complete version (Doctoral dissertation, University of Detroit Mercy, 2003). *Dissertation Abstracts International: Section B: The Sciences and Engineering, 63*, 6090.

Fabrigar, L. R., Wegener, D. T., MacCallum, R. C., & Strahan, E. J. (1999). evaluating the use of exploratory factor analysis in psychological research. *Psychological Methods, 4*, 272-299.

Gardner, H. (1983). *Frames of the mind: The theory of multiple intelligences*. New York: Basic Books.

Glass, G. V., & Hopkins, K. D. (1996). Statistical methods in education and psychology (3rd Ed.). Needham Heights, MA: Allyn and Bacon.

Glutting, J. J., Adams, W., & Sheslow, D. (2000). *Wide Range Intelligence Test*. Wilmington, DE: Wide Range Inc.

Glutting, J. J., McDermott, P. A., Watkins, M. W., Kush, J. C., and Konold, T. R. (1997). The base rate problem and its consequences for interpreting children's ability profiles. *School Psychology Review, 26*, 176-188.

Hays, J. R., Reas, D. J., & Shaw, B. J. (2002). Concurrent validity of the Wechsler Abbreviated Scale of Intelligence and the Kaufman Brief Intelligence Test among psychiatric inpatients. *Psychological Reports, 90*, 355-359.

Horn, J. L. (1985). Remodeling old models of intelligence. In B. Wolman (Ed.), *Handbook of Intelligence: Theories, Measurements, and Applications* (267-300). New York: Wiley.

Ittenbach, R. F., Esters, I. G., & Wainer, H. (1997). The history of test development. In D. Flanagan, J. Genshaft, & P. Harrison (Eds.), *Contemporary intellectual assessment: Theories, tests, and issues* (pp. 17-31). New York: Guilford.

Jensen, A. R. (1980). *Bias in mental testing*. New York: Free Press.

Jensen, J. A., & Armstrong, R. J. (1985). *Slosson Intelligence Test for Children and Adults: Expanded norms tables' application and development*. East Aurora, NY: Slosson Educational Publications.

Kamphaus, R. W. (1993). *Clinical assessment of children's intelligence: A handbook for professional practice*. Needham Heights, MA: Allyn & Bacon.

Kamphaus, R. W., Petoskey, M. D., & Morgan, A. W. (1997). A history of intelligence test interpretation. In D. Flanagan, J. Genshaft, & P. Harrison (Eds.), *Contemporary intellectual assessment: Theories, tests, and issues* (pp. 32-47). New York: Guilford.

Kaufman, A. S. (1979). *Intelligent Testing with the WISC-R*. New York: Wiley-Interscience.

Kaufman, A. S., & Kaufman, N. L. (1990). *Kaufman Brief Intelligence Test*. Circle Pines, MN: American Guidance Service.

Kaufman, A. S., & Kaufman, N. L. (2001). Time for the changing of the guard: A farewell to short forms of intelligence tests. *Journal of Psychoeducational Assessment, 19*, 245-267.

Kohs, S. (1927). *Intelligence measurement*. New York: MacMillan.

MacCallum, R. C., Widaman, K. F., Zhang, S., & Hong, S. (1999). Sample size in factor analysis. *Psychological Methods, 4*, 84-89.

Macmann, G. M., & Barnett, D. W. (1994). Structural analysis of correlated factors: Lessons from the verbal-performance dichotomy of the wechsler scales. *School Psychology Quarterly, 9*, 161-197.

McDermott, P. A. (1988). Agreement among diagnosticians or observers: Its importance and determination. *Professional School Psychology, 3*, 225-240.

McDermott, P. A., Fantuzzo, J. W., Glutting, J. J. (1990). Just say no to subtest analysis: A critique on Wechsler theory and practice. *Journal of Psychological Assessment, 8*, 290-302.

McDermott, P. A., Fantuzzo, J. W., Glutting, J. C., & Watkins, M. W. (1992). Illusions of meaning in the ipsative assessment of children's ability. *Journal of Special Education, 25*, 504-526.

Shields, J., Konold, T. R., & Glutting, J. C. (2004). Validity of the wide range intelligence test: Differential effects across race/ethnicity, gender, and education level. *Journal of Psychological Assessment, 22,* 287-303.

Silverstein, A. B. (1990). Short forms of individual intelligence tests. *Psychological Assessment: A Journal of Consulting and Clinical Psychology, 2*, 3-11.

Spearman, C. (1927). *The abilities of man: Their nature and measurement*. New York: Macmillan.

Stinnett, T. A., Havey, J. M., & Oehler-Stinnett, J. (1994). Current test usage by practicing school psychologists: A national survey. *Journal of Psychoeducational Assessment, 12*, 331-350.

The Psychological Corporation (1992). *Wechsler Individual Achievement Test*. San Antonio, TX: Author.

The Psychological Corporation (1999). *Wechsler Abbreviated Scale of Intelligence*. San Antonio, TX: Author.

Thompson, A. P. (1987). Methodological issues in the clinical evaluation of two- and four-subtest short forms of the WAIS-R. *Journal of Clinical Psychology, 43*, 142-144.

Thorndike, E. (1927). *The measurement of intelligence*. New York: Columbia University, Teachers College, Bureau of Publications.

Thorndike, R. L. (1990). Origins of intelligence and its measurement. *Journal of Psychoeducational Assessment, 8,* 223-230.

Thorndike, R. L. (1997). The early history of intelligence testing. In D. Flanagan, J. Genshaft, & P. Harrison (Eds.), *Contemporary intellectual assessment: Theories, tests, and issues* (pp. 3-16). New York: Guilford.

Watkins, M. W. (2000). *Monte Carlo PCA for Parallel Analysis* [Computer Software]. State College, PA: Author.

Wechsler, D. (1939). *The measurement of adult intelligence*. Baltimore: Williams & Wilkins.

Wechsler, D. (1991). *Wechsler Intelligence Scale for Children-Third Edition.* San Antonio, TX: The Psychological Corporation.

Wechsler, D. (1997). *Wechsler Adult Intelligence Scale-Third Edition.* San Antonio, TX: The Psychological Corporation.

Wechsler, D. (2002). *Wechsler Preschool and Primary Scale of Intelligence-Third Edition.* San Antonio, TX: The Psychological Corporation.

Wechsler, D. (2003). *Wechsler Intelligence Scale for Children-Fourth Edition.* San Antonio, TX: The Psychological Corporation.

Wilkinson, G. S. (1993). *Wide Range Achievement Test 3.* Wilmington, DE: Wide Range Inc.

Table 1

Sample Demographic Characteristics.

	Present Study		Collins (2002)	
	N	%	*N*	%
Sex				
Male	38	44.2	30	45.5
Female	48	55.8	36	54.5
Race/Ethnicity				
Caucasian	77	89.5	65	98.5
Hispanic	8	9.3	0	0
Asian	1	1.2	1	1.5
Disability				
Not Disabled	82	95.3	32	48.5
Learning Disabled (LD)	2	2.3	25	37.9
LD & Emotionally Disabled	1	1.2	0	0
Other Health Impaired	1	1.2	0	0
Mental Retardation	0	0	6	9.1
Speech/Language Disability	0	0	2	3
Developmental Delay	0	0	1	1.5
Test Type				
Volunteer	81	94.2	-	-
Psychological Evaluation	5	5.8	-	-
Grade				
1	7	8.1	-	-
2	17	19.8	-	-
3	8	9.3	-	-
4	12	14.0	-	-
5	17	19.8	-	-
6	9	10.5	-	-
7	3	3.5	-	-
8	1	1.2	-	-
9	0	0	-	-
10	2	2.3	-	-
11	1	1.2	-	-
12	3	3.5	-	-
12+	6	7.1	-	-

Note. There was no demographic information for *Test Type* or *Grade* reported in Collins (2002).

Table 2

Pearson Product-Moment Correlation Coefficients for the WASI and WRIT.

WASI	Wide Range Intelligence Test						
	Verbal	Visual	General	VOC	VA	D	MAT
VIQ	**.77 (.59) [.82]**	.51 (.26) [.67]	.75 (.56) [.80]	.75 (.56) [.78]	.59 (.35) [.73]	.37 (.14) [.67]	.53 (.28) [.56]
PIQ	.43 (.18) [.65]	**.76 (.58) [.78]**	.69 (.48) [.76]	.36 (.13) [.63]	.39 (.15) [.57]	.61 (.37) [.75]	.70 (.49) [.68]
FSIQ-4	.69 (.48) [.79]	.71 (.50) [.79]	**.82 (.67) [.85]**	.64 (.41) [.76]	.57 (.32) [.71]	.55 (.30) [.77]	.69 (.48) [.67]
VOC	.71 (.50) [.82]	.43 (.18) [.62]	.66 (.44) [.77]	**.73 (.53) [.80]**	.51 (.26) [.71]	.26 (.07) [.62]	.48 (.23) [.52]
SIM	.71 (.50) [.72]	.50 (.25) [.64]	.71 (.50) [.72]	.66 (.44) [.66]	**.58 (.34) [.68]**	.41 (.17) [.61]	.48 (.23) [.56]
BD	.36 (.13) [.59]	.64 (.41) [.67]	.59 (.35) [.67]	.33 (.11) [.55]	.30 (.09) [.54]	**.55 (.30) [.71]**	.57 (.32) [.53]
MR	.40 (.16) [.59]	.70 (.49) [.74]	.63 (.40) [.71]	.31 (.10) [.59]	.39 (.15) [.50]	.53 (.28) [.64]	**.67 (.45) [.69]**

Note. r^2s presented in parentheses. Collins (2002) correlation coefficients are presented in brackets. All correlations significant ($p < .05$). WRIT = Wide Range Intelligence Test; WASI = Wechsler Abbreviated Scale of Intelligence; VIQ = Verbal IQ; PIQ = Performance IQ; FSIQ-4 = Full Scale IQ-4 Subtests; VOC = Vocabulary Subtest; SIM = Similarities Subtest; BD = Block Design Subtest; MR = Matrix Reasoning Subtest; Verbal = Verbal IQ; Visual = Visual IQ; General = General IQ; VOC = Vocabulary Subtest; VA = Verbal Analogies Subtest; D = Diamonds Subtest; MAT = Matrices Subtest. N = 86 for present study. N = 66 for Collins (2002). Correlations in bold represent correlations between similar subtests or IQs.

Table 3

Descriptive Statistics, Dependent t-tests, and Effect Sizes for WASI and WRIT Comparisons.

			95% Confidence Interval				
	M	*SD*	Lower	Upper	*t*	*p*	Δ
WASI Verbal IQ	108.56	14.57	-1.52	2.48	0.47	.637	.03
WRIT Verbal IQ	108.08	11.43					
WASI Performance IQ	103.84	12.71	-3.51	0.21	-1.76	.081	.11
WRIT Visual IQ	105.49	12.32					
WASI Full Scale IQ-4	106.95	13.47	-2.52	0.82	-1.01	.315	.06
WRIT General IQ	107.80	11.80					
WASI Vocabulary	104.76	13.68	-5.17	-1.13	-3.10*	.003	.21
WRIT Vocabulary	107.83	11.68					
WASI Similarities	110.80	14.01	1.44	6.63	3.09*	.003	.27
WRIT Verbal Analogies	106.77	12.04					
WASI Block Design	101.99	11.84	-3.44	1.21	-0.95	.343	.07
WRIT Diamonds	103.10	11.07					
WASI Matrix Reasoning	105.15	13.55	-2.88	1.68	-0.52	.603	.04
WRIT Matrices	105.74	12.36					

Note. $df = 85$ for all comparisons. WRIT = Wide Range Intelligence Test; WASI = Wechsler Abbreviated Scale of Intelligence; Full Scale IQ-4 = Full Scale IQ-4 Subtests; Δ = Glass' Delta (Glass & Hopkins, 1996). * $p < .05$ with Bonferonni correction for family wide error rate = .007 (.05/7).

Table 4

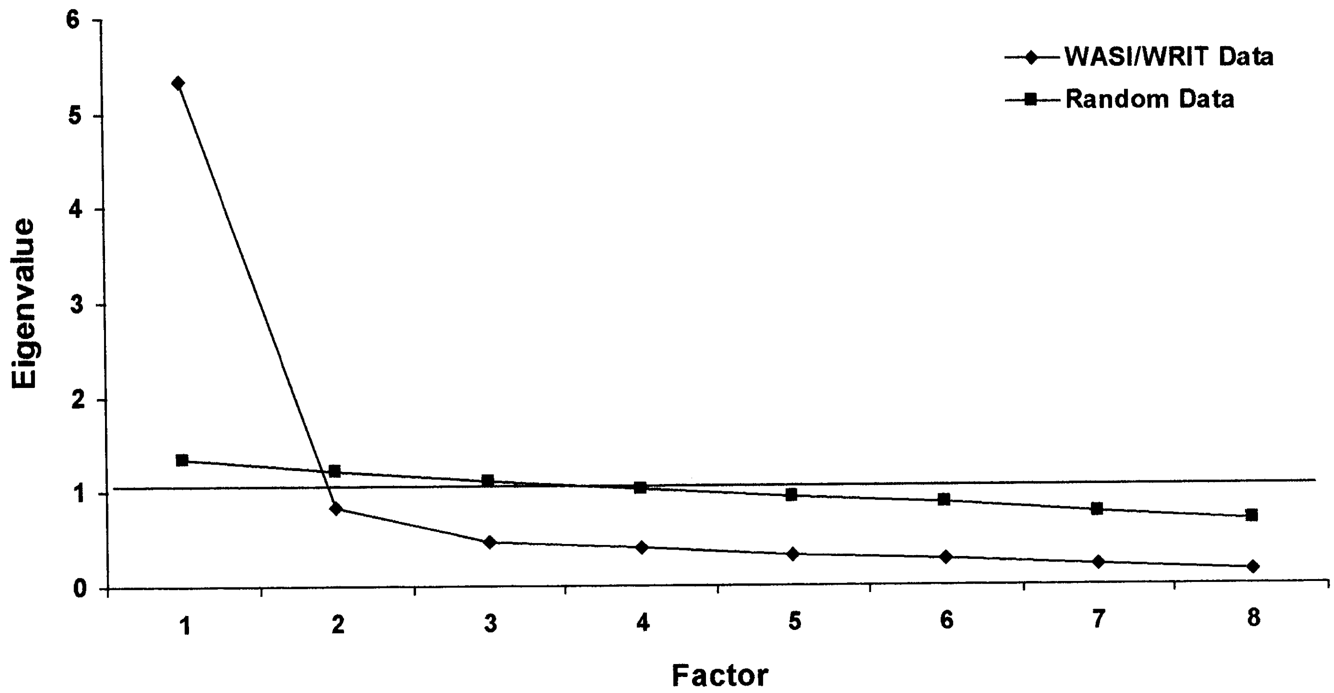
Intercorrelations and Factor Structure Coefficients for WASI/WRIT Subtest Scores.

		Correlations								Single Factor Structure Coefficient	Promax Structure Coefficient	
WASI/WRIT Subtests		WASI				WRIT					Factor I	Factor II
		Voc	Sim	BD	MR	Voc	VA	Dia	Mat	"g"		
WASI	Voc									.84	***.91***	.68
	Sim	.72								.83	***.85***	.70
	BD	.40	.37							.73	.61	***.78***
	MR	.52	.45	.59						.78	.65	***.83***
WRIT	Voc	.73	.66	.33	.31					.83	***.89***	.66
	VA	.51	.58	.30	.39	.52				.79	***.79***	.67
	Dia	.26	.41	.55	.53	.26	.30			.72	.58	***.79***
	Mat	.48	.48	.57	.67	.43	.48	.52		.78	.66	***.81***
						Eigenvalues					5.33	0.83
						Percentage of Variance					66.62	10.46
						Cumulative % of Variance					66.62	77.08

Note. N = 152. Collins (2002) sample (N = 66) and the current sample (N = 86) were merged for the factor analyses. Bold italics are salient structure coefficients that are associated with their theoretically consistent dimension. WRIT = Wide Range Intelligence Test; WASI = Wechsler Abbreviated Scale of Intelligence; WASI Voc = WASI Vocabulary, WASI Sim = WASI Similarities, WASI BD = WASI Block Design, WASI MR = WASI Matrix Reasoning, WRIT Voc = WRIT Vocabulary, WRIT VA = WRIT Verbal Analogies, WRIT Dia = WRIT Diamonds, WRIT Mat = WRIT Matrices. All Factor Structure Coefficients are ≥ .40, thus salient. Promax rotated Factor I and Factor II r = .75.

Figure Caption

Figure 1. Scree Plots for WASI/WRIT Joint Factor Structure Parallel Analysis.


6
5
4
3
2
1
0
Eigenvalue
1
2
3
4
5
6
7
8
Factor
WASI/WRIT Data
Random Data

Appendix A

TEACHER MEMO

Hello, my name is Greg Wilson, school psychology intern. I will be working under Alesia Grigg, your district's school psychologist for the 2004-2005 school year. In addition to my normal roles as a school psychologist, I will be conducting research that compares two brief measures of intelligence. I would appreciate it if you would allow me to use your students as participants. If you do not mind, please pass out these permission slips to the students in your classroom so they may take them home for their parents to sign. When permission slips are returned, please put them in my mailbox or return them to the school secretary.

Thank you for your time and effort,

Greg Wilson

Appendix B

PARENT / GUARDIAN CONSENT FORM FOR NON-CASE STUDY TESTING

Dear Parent or Guardian,

My name is Greg Wilson, a school psychology intern who will be working in your child's district this year. In addition to my normal duties as a school psychologist, I will be conducting research regarding the relationship between two brief measures of cognitive ability. It would be greatly appreciated if you would allow your child to participate in the study. Participation is completely voluntary, and no personal information will be collected. Test results will be kept confidential. However, it may be beneficial to share the results with some of your child's educators. A brief individual assessment might assist your child's teacher in planning appropriate educational strategies. Testing will be conducted during school hours at times such as study hall, P.E., and other non-core classes. In most instances the testing would take about an hour.

Your cooperation will be greatly appreciated. Thank you,

Please return this completed form to your child's teacher.

I **do** give permission for my child to participate in the study. [ ]

I **do not** give permission for my child to participate in the study. [ ]

I would like my child's results to be shared with his/her educators. [ ]

______________________	______________________	______________________
Child's Name	Parent's Signature	Date

If you have any questions regarding this study or would like to know your child's results, please contact me at (217) 348-7700.