Journal of Psychoeducational Assessment
2005, *23*, 161-172

VALIDITY OF THE COMPREHENSIVE TEST OF NONVERBAL INTELLIGENCE (CTONI)

Eric A. Rossen, Deirdre K. Shearer, Randall D. Penfield, and John H. Kranzler
University of Florida

This study examined the construct validity of the Comprehensive Test of Nonverbal Intelligence (CTONI) in two separate investigations. The first study examined criterion-related evidence of validity across racial/ethnic groups on the CTONI and the Woodcock-Johnson Psychoeducational Tests of Achievement (3rd edition). The second study examined the comparability of the standard and computerized forms of the CTONI. Results of the first study revealed that the CTONI overall score did not correlate significantly with reading achievement and correlated moderately with math achievement. The CTONI also showed significant racial/ethnic group differences, despite the absence of these differences on achievement. Results of the second study revealed that average scores on the standard form of the CTONI were significantly higher statistically than those on the computerized form. Correlations between raw scores on the computerized form of the CTONI and age were statistically significant, although generally low. In sum, results of this research raise concerns and questions about the validity of the CTONI. The CTONI should be used with caution, if at all, until further research is conducted.

Nonverbal measures of cognitive ability are playing an increasingly important role in psychoeducational assessment, based in large part on the changing demographics of American schools. Racial/ethnic minority students currently comprise approximately 40% of all students in the schools; and between 1972 and 2000, racial/ethnic minority student enrollment in public schools increased by 17%, with Hispanic students being the fastest growing student group in the nation's public schools (National Center for Educational Statistics [NCES], 2002). Moreover, in 1999, almost 20% of all 5- to 24-year-old students spoke a language other than English in the home, which is more than twice the number of students who spoke a language other than English in the home in 1979 (NCES, 2003). Some states report that as many as five dozen languages are spoken within a single school system (Unz, 1997). As the racial/ethnic and linguistic diversity of children and youth continues to increase, language- and culture-reduced measures of cognitive ability will play a larger and more central role in psychoeducational assessment than today.

To address these urgent needs, a number of nonverbal measures of cognitive ability have recently been developed, including the Universal Nonverbal Intelligence Test (UNIT; Bracken & McCallum, 1998), the Naglieri Nonverbal Ability Test (NNAT; Naglieri, 1996), the Leiter International Performance Scale-Revised (Leiter-R; Roid & Miller, 1997), the Test of Nonverbal Intelligence-Third Edition (TONI-III; Brown, Sherbenou, & Johnsen, 1997), the Reynolds Intellectual Assessment Scales (RIAS; Reynolds & Kamphaus, 2003), and the Comprehensive Test of Nonverbal Intelligence (CTONI; Hammill, Pearson, & Wiederholt, 1996). Despite the fact that these and other nonverbal tests of cognitive ability are among the most frequently used instruments with bilingual and limited-English-proficient students (Ochoa, Powell, & Robles-Pina, 1996), research on their validity has been minimal. Further research is crucial to ensure the valid interpretation and use of the results of these tests with all children and youth.

Comprehensive Test of Nonverbal Intelligence

Of these new nonverbal measures, the CTONI is one of the most widely used. The CTONI is an individually administered, standardized test of cognitive ability that can be given nonverbally, verbally, or via computer. According to Hammill et al. (1996), the CTONI measures nonverbal intelligence, which they define as "those particular abilities that exist independent of language and that increase a person's capacity to function intelligently" (p. 2). Designed for individuals between the ages of 6 and 89, the authors assert that it can be used in instances when "most other mental ability tests are either inappropriate or biased" (p. 13) and that it is appropriate for individuals who are deaf or economically disadvantaged or have neurological impairment.

The CTONI consists of six subtests that were developed to measure Analogical Reasoning, Categorical Classification, and Sequential Reasoning. Instructions can either be given verbally or in pantomime. Items are presented in a matrix format. For each item, examinees respond by pointing. If they cannot point, the examiner may point to each of the pictures and have the examinees nod when their choice is indicated. When administered by computer, examinees listen to instructions given by the computer and use a mouse to click on their answer to each item. The computer automatically scores each response and presents each item and subtest in order.

Psychometric evidence reported in the manual suggests that the CTONI's reliability is adequate. Alternate forms reliability coefficients are reported at 19 age intervals for the Pictorial Nonverbal Intelligence Quotient (PNIQ), the Geometric Nonverbal Intelligence Quotient (GNIQ), and the Nonverbal Intelligence Quotient (NIQ). Coefficient alphas for the six subtests are all above .80. The mean test-retest reliability coefficients for these scales are .93, .95, and .97, respectively. The test-retest reliability data were collected using pantomime instructions on the first testing and oral instruction on the second testing. The reliability of the computer version (CTONI-CA) is not reported, however. The lack of evidence supporting the reliability of the CTONI-CA is a significant omission that needs to be addressed.

Evidence of criterion-related validity reported in the manual suggests that the CTONI correlates with other measure of intelligence. Specifically, the CTONI was found to correlate .81 with the Wechsler Intelligence Scale for Children-Third Edition (WISC-III; Wechsler, 1991), .74 with the Peabody Picture Vocabulary Test-Revised (PPVT-R; Dunn & Dunn, 1981), and .82 with the Test of Nonverbal Intelligence-Second Edition (TONI-2; Brown et al., 1990). The PNIQ was found to correlate only .43 with the TONI-2 Quotient, however (Aylward, 1998).

In the CTONI manual, Hammill et al. provide no evidence of criterion-related validity with measures of achievement. On other nonverbal measures, correlations with achievement are substantial. For example, Bracken and McCallum (1998) reported correlations of .70 and .71 between the UNIT and the Reading and Math composites of the Wechsler Individual Achievement

Test (WIAT), respectively. Naglieri (1996) found that the NNAT correlated .56 with the SAT-9. Wiseley (2001) found that the overall score on the CTONI, the Nonverbal Intelligence Quotient (NIQ), correlated significantly with reading achievement and math achievement in a Native American population. No data substantiating the CTONI's criterion-referenced validity with achievement across other racial/ethnic groups was reported, however. Hammill et al. stated that, although economically and educationally disadvantaged ethnic/racial groups typically perform lower than average on tests of cognitive ability, the CTONI was constructed to minimize the effects of cultural, linguistic, racial, and ethnic bias, resulting in minimized score differences across groups. Data supporting these claims are not presented, however.

To examine the factor structure of the CTONI, Hammill et al. reported the results of exploratory factor analysis. They found that all of the CTONI subtests had a salient loading on a single general factor. This finding was replicated in 11 different subgroups of the school population. Thus, although the authors assert that the CTONI measures three higher-order abilities, the loadings of all items on a single factor do not support the use and interpretation of the additional scales beyond the overarching NIQ composite score. In addition, Lingen (1998) noted that tables in the manual reporting the results of classical item analysis indicate that the CTONI did not meet acceptable criteria for item discrimination and item difficulty for the youngest age groups (viz., 6-0 to 7-11 years).

In addition to these concerns about the reliability and validity evidence presented in support of the CTONI, no data are presented to support the equivalence of the pantomime, oral, and computerized versions of the CTONI. Establishing the validity of one form of the test does not ensure validity for other forms of the test, including computer-based forms (Butcher, Perry, & Atlis, 2000). Before different methods can be used for administration, empirical evidence should clearly indicate the degree to which scores are interchangeable (American Educational Research Association [AERA], American Psychological Association [APA], National Council on Measurement Education [NCME], 1999).

Purpose of the Current Research

The aim of this study is to further investigate the validity of the CTONI in two separate studies. The first study examined the CTONI's criterion-related validity with a standardized measure of achievement and compared the performance across groups of African American and Caucasian children. The aim of the second study was to examine the equivalence of the standard and computerized forms of the CTONI.

METHOD

Study 1: CTONI Construct Validity

Participants

Participants were 46 students in grades 1 through 6 (25 boys, 22 girls) from three elementary schools in north central Florida. The age of participants ranged from 6 to 10 years (M = 8.02, SD = 1.6). The racial breakdown of the sample was 27 Caucasian, 17 African American, and 2 Hispanic students. All participants were treated in accordance with the *Ethical Principles of Psychologists and Code of Conduct* (American Psychological Association, 2002).

Instruments

CTONI. The CTONI is a norm-referenced standardized test designed to measure nonverbal intellectual abilities. The test is presented in a matrix format, yielding three composite scores: the Pictorial Nonverbal Intelligence Quotient (PNIQ), the Geometric Nonverbal Intelligence Quotient (GNIQ), and the overarching Nonverbal Intelligence Quotient (NIQ). The standard battery consists of six subtests. Three subtests contribute to the PNIQ (Pictorial Analogies [PA], Pictorial Categories [PC], and Pictorial Sequences [PS]); and three subtests contribute to the GNIQ (Geometric Analogies [GA], Geometric Categories [GC], and Geometric Sequences [GS]). The NIQ is the combination of the PNIQ and GNIQ. All scaled scores have a M of 100 and a SD of 15.

Woodcock Johnson Psychoeducational Tests of Achievement, Third Edition (*WJ-III TA*). The WJ-III TA (Woodcock, McGrew, & Mather, 2001) is a norm-referenced standardized test of academic achievement. The WJ-III contains 22 subtests that assess reading, mathematics, written language, oral language, and academic knowledge. Specific combinations of the 22 subtests form composite scores. We used only those subtests that allow for determination of Broad Reading and Broad Math composite scores. The reading measures administered were Word Identification, Reading Fluency, and Passage Comprehension. The tests of math achievement administered were Addition, Subtraction, Multiplication, and Word Problems. The WJ-III TA allows for the calculation of standard scaled scores, with a M of 100 and a SD of 15.

The reliability and validity of the WJ-III TA is well established. For example, for ages 5 to 19, the Broad Reading and Math clusters have a median reliability of .93 and .97, respectively (Woodcock et al., 2001). Evidence of concurrent validity has been established with various measures of intelligence and achievement. Broad Reading and Math correlate approximately .70 and .67 with the General Intellectual Ability (GIA) score of the Woodcock Johnson Tests of Cognitive Ability. In addition, these scores correlate about .55 with the WISC-III. Furthermore, the Reading Scale on the WJ-III TA correlates .65 with the Wechsler Individual Achievement Test (WIAT) Reading Scale and .67 with the WIAT Math Scale.

Procedure

Consent forms were sent to parents/guardians in three elementary schools in north central Florida. Participants in this study were those who returned signed consent forms. Tests were administered in a counterbalanced order, with half of the students taking the CTONI first and the other half taking the WJ-III first. Students were given a short break between administrations. Total testing time was approximately 75 minutes. Two trained doctoral students conducted testing under standardized conditions.

Analysis

Means and standard deviations were calculated across racial/ethnic and gender groups. One-way ANOVAs were calculated to examine mean differences across racial/ethnic groups. Pearson product-moment correlation coefficients were calculated to examine the relationships between the CTONI and the Broad Reading and Broad Math composite scores.

RESULTS

Table 1 presents descriptive statistics for the entire sample as well as for each of the racial/ethnic groups represented. As can be seen in the table, means of the composites and scale scores on both the CTONI and WJ-III TA fall in the average range for the entire sample, as well as for each subgroup. A one-way analysis of variance (ANOVA) was conducted for each of the dependent variables (Broad Reading and Math, PNIQ, GNIQ, NIQ), with racial/ethnic group as the independent variable. Given unequal cell sizes, Levine's test was used to examine the assumption of homogeneity of variances. For each dependent variable this assumption was met. The ANOVA was significant for the NIQ and GNIQ scores, $F(2, 43) = 4.622$, $p = .015$, and $F(2, 43) = 5.78$, $p = 0.01$, respectively. However, the ANOVA was not significant for PNIQ, Broad Reading, or Broad Math scores.

Table 1

Descriptive Statistics for the CTONI and WJ-III across Groups

	CTONI					
	Total Sample[a]		Caucasian[b]		African American[c]	
Scale	*M*	*(SD)*	*M*	*(SD)*	*M*	*(SD)*
NIQ	108.4	12.6	112.7	11.9	102.6	11.7
PNIQ	105.5	13.2	108.3	13.2	101.8	13.1
GNIQ	109.8	13.0	114.7	12.8	102.6	10.4
WJ-III Reading	103.1	13.9	101.8	15.7	104.1	10.7
WJ-III Math	110.1	14.6	110.5	15.6	109.4	14.2

[a]*N* = 46. [b]*n* = 27. [c]*n* = 17.

Zero-order correlations between the composites of the CTONI and the WJ-III TA are presented in Table 2. As can be seen in this table, the CTONI scores are significantly intercorrelated. In addition, scores on the GNIQ and NIQ strongly correlated with math achievement, whereas PNIQ showed a significant but only modest correlation. The GNIQ showed a significant, albeit rather marginal correlation with reading achievement; neither the PNIQ nor NIQ correlated significantly with reading scores.

Table 2

Correlations among Composite Scores

Scale	GNIQ	PNIQ	NIQ	WJ-Reading	WJ-Math
GNIQ	-	0.61**	0.89**	0.31*	0.53**
PNIQ	-	-	0.89**	0.21	0.39**
NIQ	-	-	-	0.26	0.50**
WJ-Reading	-	-	-	-	0.57**
WJ-Math	-	-	-	-	-

*$p < .05$. **$p < .01$.

Study 2: Validity of the CTONI-CA

Participants

Participants were 34 students in grades 2 through 5 (17 boys, 17 girls) from two elementary schools in north central Florida. The age of participants ranged from 7 to 11 years (M = 8.6 years, SD = 1.5). All participants were recruited from general education classrooms. Participants were treated in accordance with the *Ethical Principles of Psychologists and Code of Conduct* (APA, 2002).

Instruments

CTONI. Both the CTONI and CTONI-CA were used. The CTONI-CA, the computerized version of the CTONI, was presented over a laptop computer with oral directions provided through the computer program. Once the examiner started the program, the examinee was asked to select a response by clicking with the mouse. The computer automatically scored each response and presented subsequent items and subtests. The examiner remained in the room with the examinee.

Procedure

Participants in this study were given both the CTONI and CTONI-CA in a counterbalanced design. Participants were randomly assigned to administration order so that half the participants (n = 17) were administered the CTONI first, followed by the CTONI-CA, and the other half of the participants in the reverse order. A mean of 3 months separated administration of the CTONI and the CTONI-CA. Two trained doctoral students conducted testing under standardized conditions.

RESULTS

Table 3 displays descriptive statistics for the composite scales and subtests by standard and computerized versions. As shown in the table, the mean scores on the CTONI were substantially higher than those on the CTONI-CA. Differences between the CTONI and CTONI-CA for NIQ, GNIQ, and PNIQ were 9.3, 11.2, and 6.0, respectively, with the CTONI yielding the higher score in each instance. In addition, *SD*s on the CTONI-CA were larger than those on the CTONI. Average *SD*s for NIQ, GNIQ, and PNIQ were 15.3 and 19.2 for the CTONI and CTONI-CA, respectively. Effect sizes were calculated by dividing the difference in means of the two formats by the mean standard deviation. For subtests, the effect sizes ranged from a small effect for GC (κ= 0.2) to a large effect for the PC (κ= 0.8), with an overall mean effect size of 0.44. Effect sizes for the composite scales were in the medium effect range. Although all means are in the average range, the effect sizes are surprisingly large for what are supposedly *parallel forms of the same test.*

Table 3
Descriptive Statistics on the CTONI and CTONI-CA

	CTONI		CTONI-CA		
Scale	*M*	*(SD)*	*M*	*(SD)*	κ
NIQ	106.2	15.2	96.9	19.1	0.5
PNIQ	105.1	16.1	93.9	20.1	0.6
GNIQ	106.7	14.7	100.7	18.5	0.4
PA	9.9	3.5	8.9	3.2	0.3
GA	10.0	2.9	8.9	3.7	0.4
PC	11.5	2.3	9.0	4.1	0.8
GC	12.1	3.2	11.6	3.5	0.2
PS	11.1	3.5	9.2	4.7	0.5
GS	10.9	2.6	9.9	3.6	0.3

N = 34.

Table 4
Analysis of Variance of Format, Order, and Interaction Effects

Scale	Order	Format	Interaction
NIQ	$F = 4.51$*	$F = 25.91$**	$F = 2.22$
	$p = .04$	$p = .00$	$p = .15$
PNIQ	$F = 2.43$	$F = 24.13$**	$F = 0.97$
	$p = .13$	$p = .00$	$p = .33$
GNIQ	$F = 6.38$*	$F = 7.616$**	$F = 1.33$
	$p = .02$	$p = .01$	$p = .26$
PA	$F = 1.84$	$F = 5.73$*	$F = 0.01$
	$p = .19$	$p = .02$	$p = .94$
GA	$F = 3.45$	$F = 4.49$*	$F = 1.90$
	$p = .07$	$p = .04$	$p = .18$
PC	$F = 0.78$	$F = 14.66$**	$F = 0.39$
	$p = .38$	$p = .00$	$p = .54$
GC	$F = 3.04$	$F = 0.89$	$F = 0.10$
	$p = .09$	$p = .35$	$p = .76$
PS	$F = 2.19$	$F = 11.70$**	$F = 1.30$
	$p = .15$	$p = .00$	$p = .26$
GS	$F = 5.77$*	$F = 5.84$*	$F = 0.58$
	$p = .02$	$p = .02$	$p = .45$

*$p < .05$. **$p < .01$.

Table 4 displays the results of a split-plot ANOVA conducted to determine the effects of the administration format, the order of administration, and the interaction between administration format and order. An α of .05 was used for all statistical tests. As shown in this table, administration format was significantly related to all composite scales and subtests, with the exception of GC. The ANOVA also indicates significant order effects on composite scales of NIQ and GNIQ as well as for the GS subtest. The ANOVA results further demonstrate significant mean differences between versions. No interaction was found between administration format and order.

Table 5 displays correlations between scores obtained on the CTONI and the CTONI-CA. As can be seen in this table, correlations among the NIQ, GNIQ, and PNIQ are substantial and statistically significant. Correlations ranged from .39 to .83 ($M = .66$). However, correlations among subscales were surprisingly low across what are supposedly parallel forms of the same test, and in one case, for PC, the correlation was not statistically significant.

Table 5
Correlation between CTONI Administration Formats

Scale	*r*
NIQ	.83*
PNIQ	.72*
GNIQ	.75*
PA	.76*
GA	.52*
PC	.39
GC	.54*
PS	.74*
GS	.73*

*$r > .50$ at .05.

Table 6
Correlation between Raw Score on the CTONI and CTONI-CA and Age (in Months)

Scale	*r*	*r*
	CTONI	CTONI-CA
GNIQ	.61*	.60*
PNIQ	.65*	.60*
PA	.63*	.62*
GA	.49	.61*
PC	.49	.29
GC	.55*	.34
PS	.51*	.53*
GS	.51*	.56*

*$r > .50$ at .05.

Table 6 displays the correlation between raw score and age for both the CTONI ($M = .49$) and the CTONI-CA ($M = .46$). Correlations between mental test scores and age are expected to be significant, demonstrating the ability of the measure to differentiate between different age groups. As can be seen in this table, scores on the PC and GC subtests were not statistically significantly correlated with age when administered over the computer. In addition, correlation coefficients for subtests GA and PC on the CTONI were not significantly correlated with age.

DISCUSSION

The purpose of this research was to examine the validity of the CTONI, a widely used nonverbal measure of cognitive ability. More specifically, we examined (a) the CTONI's relationship to a standardized measure of academic achievement and performance across racial/ethnic groups and (b) the comparability of the computerized version of the CTONI with the standard version.

Results of the first study revealed that GNIQ of the CTONI correlated moderately, yet statistically significantly, with math and reading achievement. However, correlations for PNIQ and NIQ did not correlate significantly with reading, and correlations with math were only .39 and .50 for the entire sample. In addition, the CTONI showed significant racial/ethnic group differences, despite the absence of racial/ethnic group differences on achievement. It would be expected that for a population with equivalent achievement scores across groups, the CTONI scores would also be consistent across groups given that children are often expected to achieve at a level comparable with IQ, because tests of intellectual ability are often used as benchmarks for achievement. Therefore, one would assume some consistency of results across IQ and achievement tests. This finding does not support Hammill et al.'s (1996) contention that this test will minimize these group differences.

Problems were also noted in the pattern of correlations between the CTONI and external criteria of validity. All CTONI composites correlated significantly with WJ-III TA Broad Math. There was also a significant correlation between the CTONI PNIQ and WJ-III TA Broad Reading. However, no statistically significant relationship was found between scores on the GNIQ and NIQ on the CTONI and the WJ-III TA Broad Reading. The absence of a significant correlation between these measures raises questions concerning the criterion-related validity of the CTONI. Perhaps most troubling is the NIQ's low correlation with reading.

The second study compared scores obtained from the computerized and standard forms of the CTONI. Since its publication, the CTONI-CA has been available for commercial use even though virtually no information has been reported regarding its validity, reliability, or equivalency with the standard administration. To support the validity and subsequent use of the CTONI-CA, correlations should be high enough to support a claim of parallel forms. Results of this study, however, raise numerous concerns regarding the comparability of the CTONI and CTONI-CA. The standard form of the CTONI was found to yield, on average, a statistically significantly higher score than the computerized form. Correlations between the two forms were also generally quite low, ranging from .39 to .83 with a mean of .66, suggesting that scores for these forms are not interchangeable. Given the high test-retest reliabilities reported by Hammill et al. (1996) for the CTONI, the mean interval of 3 months between administration of the two forms of the test is an implausible explanation of the observed differences. Similarly problematic are the correlations between raw scores on the CTONI-CA and age. In addition to the finding that scores on several subtests did not correlate significantly with age, the overall correlations are rather low for a test of cognitive ability, suggesting that the CTONI-CA does a marginal job of discriminating across age groups (cf. Lingen, 1998).

As noted above, Hammill et al. (1996) do not indicate in the CTONI manual whether the CTONI was standardized using pantomime, oral, or computerized directions. As Kline (2000) has stated, "If the computer version is to be regarded as identical with the standard version, the correlation between the two formats should be at least 0.9" (p. 97). Results of this study, in contrast, revealed a correlation of only .66 between the two forms of the CTONI. This finding indicates that the scores on the CTONI and CTONI-CA are not equivalent.

Differences in mean scores may result from the CTONI-CA administration procedures. When administered in the computerized format, if examinees respond incorrectly to any one of three practice items they automatically receive a score of 0 for the subtest and are moved on to the next subtest. Also, the fact that examinees are unable to change an answer once clicked with the mouse may contribute to lower scores on the computer format. If an examinee accidentally clicks an answer and immediately wishes to change it, there is no way to go back, unlike when the item is administered in the standard format. In this respect, impulsive children are not given the opportunity to self-correct. Another possible explanation for differences across formats is that the program used to present items is flawed in some way, not the items themselves. In such a case, the computer program introduces construct-irrelevant variance, impacting the scores received. Further research is needed to answer these questions.

Implications for Psychologists

The assessment of racial/ethnic and linguistically diverse populations is an area of increasing concern in today's society. The introduction of new nonverbal tests of cognitive ability is an area of great importance in psychoeducational assessment. At the same time, these new measures must be reliable and valid for their intended purposes. With regard to one of these measures, the CTONI, results of this research indicate serious concerns regarding its validity and use as a benchmark for academic achievement in the schools. Not only were correlations with achievement marginal at best, but the pattern of mean scores across tests also differed for Caucasians and African Americans. Moreover, results of this research raise concerns about the comparability of scores across different forms of the CTONI. In sum, based on our findings, the CTONI and CTONI-CA should be used with caution, if at all.

Limitations

One potential limitation of this study was the sample size. Studies 1 and 2 consisted of 46 and 34 participants, respectively. Small samples are considered a limitation due to a potential lack of power. However, in the current studies there was more than enough power to reject the null hypotheses in both studies.

Future Research

Additional independent analyses of the CTONI's validity are needed. For example, confirmatory factor analysis would shed considerable light on the test's structure. The test authors claim that the test has one factor, which they call nonverbal intelligence. However, the theoretical structure of the test is hierarchical. In addition, comparisons with other standard measures of intelligence (e.g., WISC-IV) are greatly needed. For example, a joint confirmatory factor analysis with a standard measure such as the WISC-IV or WJ-III would help clarify the constructs being measured by the CTONI. Information is also needed on how the CTONI predicts performance of students across racial/ethnic groups. Studies that include students with disabilities are also needed to demonstrate that this measure is an adequate measure for assessing diverse populations. Future studies with the CTONI-CA should be geared toward assessing age differences in test taking over the computer as well as concurrent levels of anxiety. Similarly, further research on the component of exposure to computers may guide examiners on the utility of such a measure.

REFERENCES

American Educational Research Association, American Psychological Association, & National Council on Measurement Education. (1999). *Standards for educational and psychological testing*. Washington, DC: American Educational Research Association.

American Psychological Association. (2002). *Ethical principles of psychologists and code of conduct*. Washington, DC: Author.

Aylward, G. P. (1998). Review of the Comprehensive Test of Nonverbal Intelligence. In J. C. Impara & B. S. Plake (Eds.), *Mental measurements yearbook* (13th ed., pp. 310-312). Lincoln, NE: University of Nebraska Press.

Bracken, B. A., & McCallum, R. S. (1998). *Universal Nonverbal Intelligence Test* (UNIT). Itasca, IL: Riverside.

Brown, L., Sherbenou, R. J., & Johnsen, S. K. (1990). *Test of Nonverbal Intelligence-Second Edition*. Austin, TX: PRO-ED.

Brown, L., Sherbenou, R. J., & Johnsen, S. K. (1997). *Test of Nonverbal Intelligence-Third Edition*. Austin, TX: PRO-ED.

Butcher, J. N., Perry, J. N., & Atlis, M. M. (2000). Validity and utility of computer-based test interpretation. *Psychological Assessment, 12*, 6-18.

Dunn, L. M., & Dunn, L. M. (1981). *Peabody Picture Vocabulary Test-Revised*. Circle Pines, MN: American Guidance Service.

Hammill, D. D., Pearson, N. A., & Widerholdt, J. L. (1996). *Comprehensive Test of Nonverbal Intelligence*. Austin, TX: PRO-ED.

Kline, P. (2000). *Handbook of psychological testing* (2nd ed.). New York: Routledge.

Lingen, G. V. (1998). Review of the Comprehensive Test of Nonverbal Intelligence. In J. C. Impara & B. S. Plake (Eds.), *Mental measurements yearbook* (13th ed., pp. 312-314). Lincoln, NE: The University of Nebraska Press.

Naglieri, J. A. (1996). *Naglieri Nonverbal Ability Test*. San Antonio, TX: The Psychological Corporation.

National Center for Education Statistics. (2002). *Elementary/Secondary education: Racial/Ethnic distribution of public school students*. Retrieved April 26, 2004, from http://nces.ed.gov/programs/coe/2002/section1/indicator03.asp.

National Center for Education Statistics. (2003). *Elementary/Secondary education: Language minority students*. Retrieved April 26, 2004, from http://nces.ed.gov/programs/coe/2003/section1-

/indicator04.asp.

Ochoa, S. H., Powell, M. P., & Robles-Pina, R. (1996). School psychologists' assessment practices with bilingual and limited-English proficient students. *Journal of Psychoeducational Assessment, 14*, 250-275.

Reynolds, C., & Kamphaus, R. (2003). *Reynolds Intellectual Assessment Scales.* Austin, TX: PRO-ED.

Roid, G. H., & Miller, L. J. (1997). *Leiter International Performance Scale-Revised.* Oakdale, IL: Stoelting.

Unz, R. (1997, October 19). Perspective on education; bilingual is a damaging myth; a system that ensures failure is kept alive by the flow of federal dollars. A 1998 initiative would bring change. *Los Angeles Times*, Opinion section, part M, p. 5.

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

Wiseley, M. (2001). Nonverbal intelligence and Native American Navajo children: A comparison between the CTONI and the WISC-III. *Dissertation Abstracts International Section A,* 62(6-A), 2030.

Woodcock, R. W., McGrew, K. S., & Mather, N. (2001). *Woodcock-Johnson III.* Itasca, IL: Riverside.